北京智源人工智能研究院
BEIJING ACADEMY OF ARTIFICIAL INTELLIGENCE

2020 年

人工智能的认知神经基础白皮书

Brain and Machine Intelligence



2021 年 3 月

人工智能的认知神经基础重大研究方向

VISIT…

LANZAROTE Caliente.com

版权声明

该《人工智能的认知神经基础白皮书》（2020 年）由北京智源人工智能研究院“人工智能的认知神经基础”方向所著，旨在帮助相关领域从业者梳理认知神经科学领域的最新进展、洞察发展趋势，并为人工智能及相关交叉学科的发展提供参考。本白皮书著作权受法律保护，转载、摘编、翻译或利用其他方式使用本白皮书观点的，应注明来源。

指导老师

刘　嘉　智源首席科学家，清华大学心理学系教授
方　方　智源研究员，北京大学心理与认知科学学院教授
陈良怡　智源研究员，北京大学未来技术学院教授
吴　思　智源研究员，北京大学信息科学技术学院教授
宋　森　智源研究员，清华大学生物医学工程系研究员
余　山　智源研究员，中国科学院自动化研究所研究员

编写组成员

徐琳璐　智源博士后，人工智能的认知神经基础重大研究方向
陈路瑶　智源博士后，人工智能的认知神经基础重大研究方向
刘　祥　智源博士后，人工智能的认知神经基础重大研究方向
邹晓龙　智源博士后，人工智能的认知神经基础重大研究方向

前 言

人工智能的终极目标是研究使计算机来模拟人的思维过程和智能行为(如学习、推理、思考、规划等)，制造类似于人脑智能的机器。为了实现这个目标，理解构成智能必要的元素，必须研究人和动物智能的载体生物大脑的本身结构以及其运行方式。神经科学聚焦于大脑如何进行信息加工，力图阐明复杂行为之下的机制和物理实现形式;认知科学研究研究人类的心智和认知过程,通过构建认知框架,预测复杂行为,揭示智能的产生机制。认知科学、神经科学与信息学科分别从不同的路径来探索智能本质的，它们的交叉碰撞出炫丽的火花。北京智源人工智能研究院的人工智能的认知神经基础重大研究方向,期望利用本方向的多学科交叉基础，用白皮书的形式总结过去一年间在这个前沿领域的一些重要探索和发现，更好的为这几个方向的研究者搭建沟通的平台和桥梁。

解析大脑和神经系统的连接方式及工作原理离不开借助先进的技术,因此高时空分辨率的生物荧光成像是系统神经科学的必要工具。康奈尔大学的 Chris Xu 研究组，一直研究三光子显微成像，实现成年斑马鱼全脑任意位置的深层高分辨率成像，相关技术也被应用于大型的模式动物如猕猴的深层脑区。高分辨率显微镜的另一个缺点是往往只能够成像很小的区域，无法看到多个神经回路和脑区的活动。为了克服这个问题,清华大学戴琼海团队一直在发展大视场、可以观察小鼠全脑很大视场内结构和功能变化的单光子显微镜。为了提高单光子成像的对比度和实现多层同时成像，他们开发了一种新的多平面宽场荧光显微镜，同时具备光学层切、体成像能力以及 1.56 mm×1.83 mm 的超大视场，并用于稀疏的小鼠血管和斑马鱼神经元成像。单个神经元脉冲发生的时间尺度在毫秒量级，因此如何提高成像的时间分辨率、用成像来观察动作电位的发生和传播也是领域内努力的重点。加州大学伯克利分校的 Na Ji 团队开发全光无源激光扫描高速成像，实现 XY 方向上的 3000 Hz 扫描的双光子成像；发展基于轴向扩展的贝塞尔光扫描双光子显微镜，实现稀疏标记样本的每秒 99 帧的体成像速率。哥伦比亚大学的 Hillman 研究组发明了可以用于小鼠等哺乳动物的单镜头的单光子光片显微镜，实现高时空分辨率的体成像。最后，为了解决如何在自由行为动物上实现高分辨率成像，北京大学分子医学研究所牵头开发了超高时空分辨微型化双光子在体显微成像系统,实现了小鼠自由行为过程中大脑神经元和神经突触活动清晰的图像,并赋予微型化系统的体成像和多平面成像能力。伴随着不断发展出来的遗传编码或是化学合成的特异性荧光探针,高时空分辨率成像能够观察越来越多的神经递质和神经肽的动力学释放过程,从而描绘出正常和神经疾病时候它们调控神经元、回路、脑区活动以及整体动物行为的变化。

这些新工具的出现所揭示的新发现往往会颠覆人们对于不同层次上神经活动发生机制的传统认知。例如，突触水平上，马普佛罗里达神经研究所的 Fitzpatrick 研究组通过光电联合实验，研究雪貂的初级视皮层神经元发现，不同刺激激活的全部突触（包括强、弱突触）共同“投票”所决定其神经元选择特异性,挑战了赫布理论所认为的强突触调控神经元选择特异性的决定性作用。神经元水平，Allen Brain Observatory 的 Christof Koch 和 Clay Reid 研究组通过分析大规模标准化神经活动数据，揭示小鼠视皮层中存在大量低特异性神经元，表明之前的视觉神经元-特定视野空间-特定特征的视觉信息模型的局限性。在视觉系统上，同样是艾伦研究所的 Christof Koch 和 Shawn R. Olsen 研究组，通过构建新型的神经电信号记录系统和神经像素电极阵列（Neuropixels），成功的观察到了小鼠视觉系统，存在如灵长类一样的功能层级活动。全脑水平上，Alipasha Vazir 研究组通过高通量全脑成像辅助预测斑马鱼决策行为，将这个行为划分成 6 个具有不同功能特征和解剖区域定位的神经元类群，并发现了小脑中的颗粒细胞神经活动对于运动的决策预测是十分重要。除了观察，神经调控手段的出现也可

以实现个体化神经认知相关的调控,如调节非侵入式的经颅电刺激等用于增强视觉运动学习和减弱视觉拥挤效应等以及新型侵入式神经调控技术手段用于重建大脑视觉认知功能等。

由于神经科学中的新技术的发展正在产生海量的数据,大脑与人工智能算法的交互变得越来越重要。目前深度神经网络的一大问题是其对噪音的鲁棒性不够。2019 年在 Nature 发表了的小鼠视觉系统的大规模的神经记录开始阐明了其群编码原则,发现其在高效和保持泛化性之间做了一个权衡。利用了该原理的人工神经网络具有更好的鲁棒性。另外 MIT 的 Joel Dapello 等人发现在传统神经网络前面加入一个模拟大脑初级视觉皮层的模型也可以提高其鲁棒性。借鉴小动物的神经环路设计也许可以帮助我们设计更高效的人工智能系统。2020 年在 Nature Machine Intelligence 上发表的 MIT 相关论文借鉴了秀丽线虫的神经系统中神经元的动态特性和连接的稀疏性和反馈结构,在跟踪道路的自动驾驶任务中比传统神经网络具有更好的鲁棒性和可解释性。

人工智能系统的进展也为帮助理解大脑的计算原理提供帮助。大脑多巴胺神经元和我们的情绪和动机息息相关，当结果好于预期时，多巴胺会被更多的释放，带来愉快的心情。但对未来的预期经常是不确定的，用一个概率分布来表示更合理，这种被称为分布式强化学习的算法最近几年取得了很好的效果。Will Dabney 等人在 2020 年的 Nature 上发文指出，大脑中不同多巴胺神经元表征的值也形成一个分布，和分布式强化学习算法的预测吻合。目前人工智能的一个前沿是对抽象关系的推理，大脑中的海马和内嗅皮层系统和空间导航息息相关，最近被证明在表征抽象关系中也发挥重要作用。在 2020 年的 Cell 文章中，Whittington 等借鉴贝叶斯学习等人工智能方法建立的模型为该系统中多种类型神经元表征提供一个统一的解释。另外，目前支持深度学习成功的反向传播算法无法在大脑中直接实现,在大脑中寻找能够完成类似功能的机制也是一个活跃的研究方向。

在当下人工智能的浪潮中，人类社会似乎从来没有像如今这样对“研究心灵本质”这个古老的问题产生过如此浓厚的兴趣，但力图从不同的方向来探寻这个问题答案的计算、神经、认知这三门学科本身存在着巨大的学科差异和鸿沟。Stephen Hawking 曾经说过，“Intelligence is the ability to adapt to change”。正确承认及认识学科间差异，在多学科的交叉区域相互沟通交流借鉴，可以从研究技术、揭示机制以及计算原理等方面来相互启发，促进不同学科的螺旋式相生相长。帮助研究人员更好的理解大脑工作的硬件、算法和实现，了解神经疾病背后机理并找到治疗方案，真正实现终极目标：“know yourself”。

目 录

前 言 1
第一部分 认知与神经科学研究进展 4
1.1 神经递质：多巴胺神经元与瞬时遗忘 4
1.2 神经元突触：多突触输入调控神经元选择特异性 5
1.3 视觉神经元：小鼠视皮层中存在大量低特异性神经元 7
1.4 视觉系统：小鼠视觉系统的功能层级 8
1.5 全脑结构连接组：全连接组测序快速构建小鼠全脑连接组图谱 9
1.6 全脑功能连接组：高通量全脑成像辅助预测斑马鱼决策行为 10
1.7 认知理论: 新证据引发修订传统认知理论的思考 12
1.8 总结与展望 14
第二部分 新技术在认知与神经科学领域的应用 16
2.1 成像技术在神经科学领域的应用 16
2.1.1 多光子深层成像技术 16
2.1.2 超大视场成像 18
2.1.3 超高速成像 19
2.1.4 自由行为动物高分辨率成像 21
2.2 探针技术在神经科学领域的应用 22
2.2.1 化学合成的荧光探针 22
2.2.2 遗传编码的荧光探针 22
2.3 神经调控技术在认知科学领域的应用 23
2.3.1 个体化非侵入式神经调控技术应用于大脑视觉认知功能的增强 23
2.3.2 新型侵入式神经调控技术应用于大脑视觉认知功能的重建 24
2.4 总结与展望 32
第三部分 神经科学对人工智能的启发 35
3.1 秀丽线虫的神经网络用于运动控制 35
3.2 果蝇的神经网络用于自然语言处理 36
3.3 模拟生物视觉系统提高对噪声的鲁棒性 36
3.4 模拟神经群编码机制提高对噪声的鲁棒性 38
3.5 总结与展望 39
第四部分 人工智能对神经科学的启发 41
4.1 强化学习理论帮助理解大脑的奖惩机制 42
4.2 反向传播算法在大脑可能实现的机制 43
4.3 神经网络建模帮助理解大脑海马中信息表征的机制 45
4.4 总结与展望 46
第五部分 结语 49
关于我们 51

第一部分 认知与神经科学研究进展

无论简单的知觉探测，还是复杂的认知决策，都与我们神经系统的结构基础与电化学活动密不可分。过去的一年中，随着多种新兴技术的发明和发展，认知与神经科学领域从微观尺度观测到宏观理论修订都取得了重大突破。一方面，在个体神经元尺度的研究向着更为精细的观测角度拓展，光电联合等技术手段让我们得以观测到更为丰富与详细的结构与活动信息；另一方面，从生物体全脑尺度，实施更多神经元的电生理记录、跨更多脑区的信号采集、或者开发更大的成像视野等技术让我们得以同时观测大脑中所有神经元进行同一任务时如何协同计算；此外，人类在体高场强磁共振、颅内单神经元记录等高精度测量技术的发展，让我们得以对社会认知、记忆、概念等高级认知功能的神经基础一探究竟。接下来，本章节将介绍在 2020 年中随着新技术的涌现，在神经活动的不同层次（递质—突触—神经元—视觉系统—全脑—人类高级认知）取得的一些重大突破。

1.1 神经递质：多巴胺神经元与瞬时遗忘

在以往对生物学习记忆的研究中，往往容易忽略一个重要的环节，那便是遗忘。近年来，随着科学家对遗忘研究的逐渐关注，人们发现遗忘并不是记忆的简单消失，而是一种大脑主动发起的生物过程，是大脑记忆管理系统的重要组成部分。遗忘的产生对生物脑的正常运行是十分重要的。与生物的学习记忆过程相比较，遗忘正是人工神经网络缺失的一环。那么，既然遗忘对于生物神经加工具有十分重要的意义，研究遗忘的神经机制，以及其在生物学习记忆过程中的角色和重要意义，也许会对人工神经网络的优化提供不同的见解。

遗忘包括永久性遗忘（永久丢失先前的记忆）和瞬时性遗忘（记忆仍然储存在脑中，但是处于一种暂时性检索受损的状态）。2021 年 1 月由 John Martin Sabandal 等人发表在 Nature 杂志上的工作[6]，进一步推进了对瞬时遗忘神经机制的研究。这项研究是以神经系统相对简单，并已得到广泛研究的果蝇为研究对象进行的。在实验中，研究者向果蝇呈现两种气味，并在其中一种气味呈现时向果蝇施加轻微的电刺激，以此来进行厌恶条件化训练。经此训练后，这一气味对于果蝇而言便会成为一种厌恶刺激，它们会条件反射性的避开这一气味。而另外一种气味因没有进行厌恶条件化，便作为中性刺激存在。随后将果蝇放入一个一端存有厌恶气味，另一端存有中性气味的 T 迷宫中进行记忆检测，发现经过训练的果蝇更偏好于有中性气味的一端。这一现象表明，果蝇已经学会了这一任务，并产生了相关记忆。为了研究瞬时遗忘，研究者会在经过厌恶条件化训练的果蝇开始进行记忆检测前，向其施加一个短暂的干扰刺激，如向果蝇吹气、电击或者施加一个蓝光刺激。以此来检测这些干扰刺激是否会对果蝇的记忆提取产生影响。类似的研究方法在人类的心理物理实验中经常会使用到。实验发现，这些短暂的干扰刺激确实会对先前的学习任务的记忆提取产生影响。经历过干扰的果蝇，对厌恶气味不再表现出明显的躲避行为。但是之所以认为干扰刺激是对记忆的提取产生了影响，是因为对先前学习任务的记忆并没有完全消失，而是出现了暂时性的遗忘，在干扰刺激施加一个小时后，经过厌恶条件化训练的果蝇再次表现出对厌恶气味的躲避行为。

以这样一个巧妙的瞬时遗忘行为范式为基础，John Martin Sabandal 等人对瞬时遗忘的神经机制展开了进一步的研究。它们发现，这一瞬时遗忘现象的产生，与果蝇的 PPL1（protocerebral posterior lateral 1，含有 12 个多巴胺神经元）多巴胺神经元有关，尤其是 PPL1 中与α2α’2 蘑菇体神经元对应的多巴胺神经元。PPL1 中多巴胺神经元的激活，能够有效抑制果蝇对先前记忆的提取；而对这些神经元的抑制，则会消除干扰刺激对果蝇躲避厌恶气味记忆的干扰。并且通过 RNA 干扰技术，研究者发现α2α’2 中的 DAMB 受体也参与其中。

DAMB 受体的表达降低后，果蝇的瞬时遗忘现象也随之消失。由此可见多巴胺在调节瞬时遗忘与记忆提取过程中起到了至关重要的作用。而在以往的研究中，人们认为多巴胺与学习记忆密切相关。这一神经递质被认为参与到学习记忆过程中相关环路突触的可塑性。可见，为了更全面的了解生物神经系统的工作机制，除了神经元间的组织结构连接、不同神经元在加工不同信息时的响应特性外，神经递质在智能过程中的作用，以及相关的工作机制也是至关重要的。

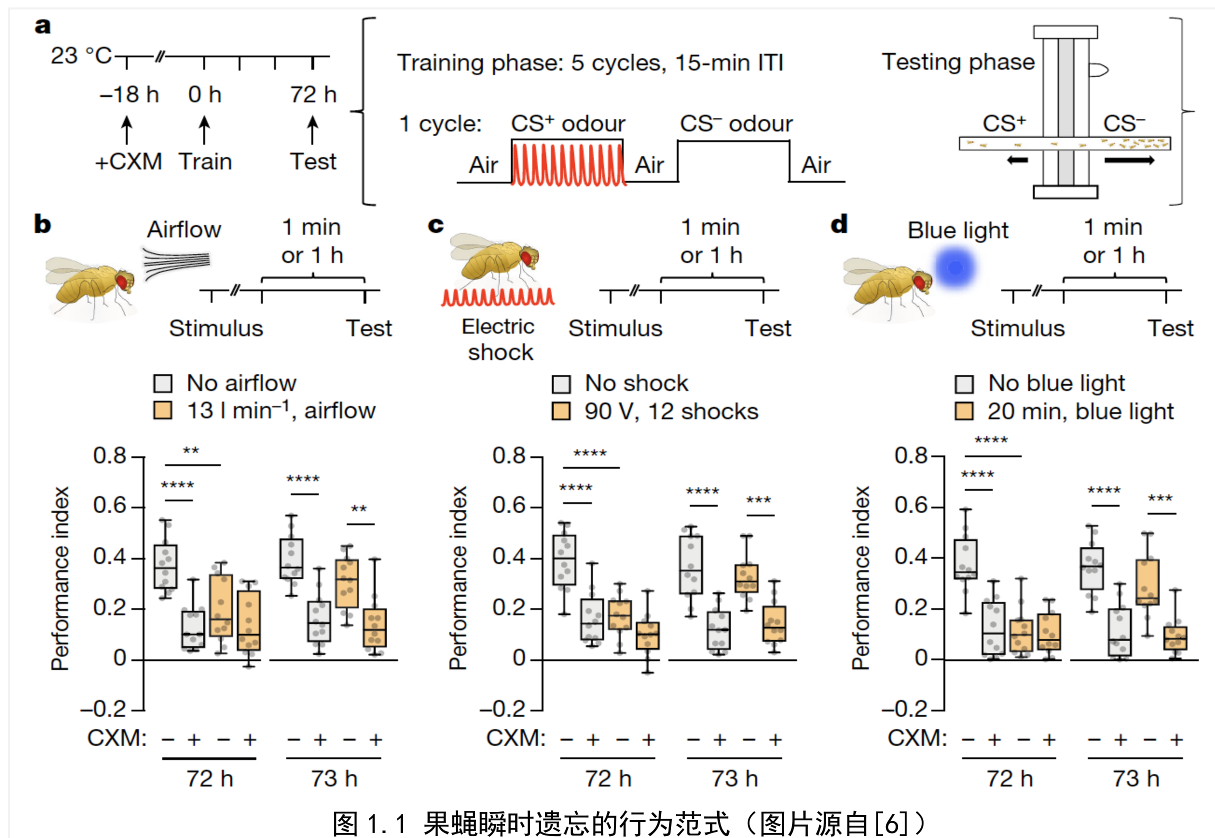


图 1.1 果蝇瞬时遗忘的行为范式（图片源自[6]）

a)厌恶气味条件化行为范式示意图。b)-d)果蝇在接受干扰刺激（吹风(b)，电击(c)，蓝光(d)）时行为范式示意图，及其分别在干扰刺激呈现后（72h）和干扰刺激呈现后一小时后（73h），进行记忆检测时的行为表现。

1.2 神经元突触：多突触输入调控神经元选择特异性

近年来光电联合技术的应用，为神经科学家提供了更为丰富的信息。结合光学成像与电镜成像技术，神经科学家能够同时获得具有超精细组织解剖结构和神经动态响应活动信息的数据，从而将神经元间的“硬件连接”与功能响应特性整合在一起。

在大脑新皮层中，单一神经元通过突触接收源自不同神经元的大量信号输入，并被这些兴奋性输入所驱动，构成了神经元对感觉输入不同特征的选择特异性基础。一直以来，人们认为这些神经元间的兴奋性连接会在发育过程中逐步成熟，通过兴奋依赖的赫布可塑性原则，选择性增强那些同步活动的突触连接，并削弱另一些不同步的突触连接。神经元上有一些突触相对更强（强突触），一些突触相对较弱（弱突触），赫布可塑性原则提示，，这些强突触驱动了神经元的选择特异性而弱突触则与神经元的信号输出以及兴奋调节的关系较弱。2020 年 12 月，Benjamin Scholl 等人通过光电联合实验，向这一假设提出了挑战。该工作发表在 Nature 杂

志上[1]。

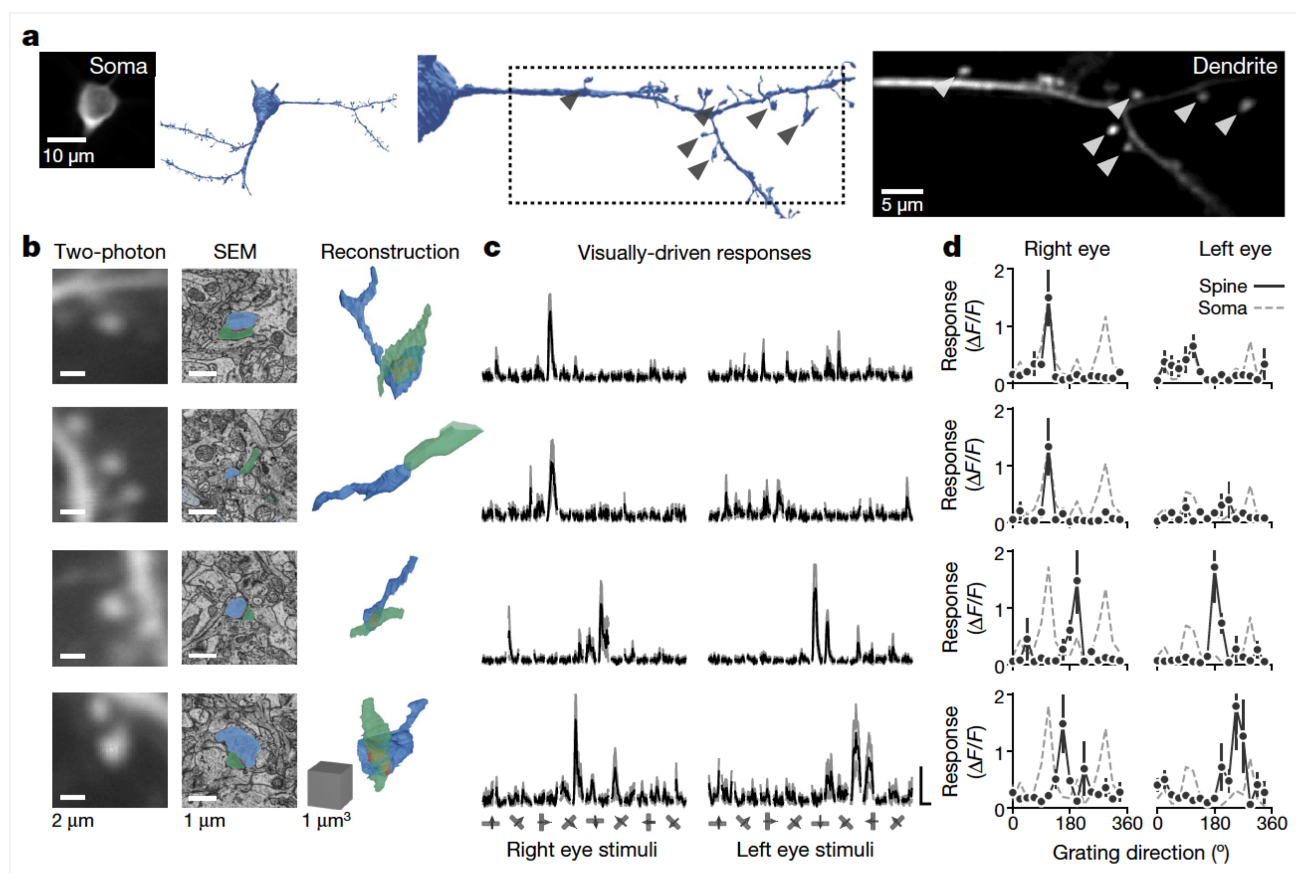


图 1.2 在体突触光学成像与扫描电镜成像的关联展示（图片源自[1]）

a)示例神经元的双光子成像效果与电镜重构效果展示图。图左和图右分别是神经元胞体和其树突双光子成像的平均投影。中间两幅图展示了这一神经元的电镜重构结果。箭头指出了对视觉刺激有响应的突触，及其在两种结构图中的定位。b)示例突触的双光子成像平均投影（左）、单层电镜成像图（中），和电镜重构效果（右）。c)图 b 中所示的突触被视觉刺激诱发的神经响应的钙信号。d)图 b 中所示突触及其对应的神经元胞体被视觉刺激诱发响应活动（ΔF/F）的对比。有些突触的响应与胞体相似，有些则不同。

在该实验中，Benjamin Scholl 等人首先通过在体双光子成像技术，采集了雪貂初级视皮层中单一椎体神经元及其树突棘的神经活动信号。随后将包含这一神经元的组织进行灌注固定，并使用 SBF-SEM（serial block-face scanning electron microscopy）技术对这一神经元及其树突棘的解剖结构进行高分辨率的扫描电镜成像。经过重构，可以获得该神经元胞体、树突，以及其上每一个突触的头部、颈部、突触后密集区（PSD，postsynaptic density）和突触前终扣（presynaptic bouton）信息。与双光子获得的影像学数据相比，电镜成像获得的超微结构更加复杂多样，对单个突触的重构所获得的结构特征可作为这一突触强度的度量。结合视觉刺激诱发的神经功能信号，该研究未发现支持初级视皮层神经元选择特异性由强突触驱动的证据。相反，他们认为单一神经元的选择特异性，似乎是由被不同刺激激活的全部突触（包括强、弱突触）所共同“投票”所决定的。强突触的响应活动是比较独立的，而弱突触往往在空间上相邻，成簇聚集在一起，而且功能比较一致。因此弱突触可以通过时间上共同响应以及空间上的聚类激活，增强共同组成的‘联合突触’的强度，以此来提高弱突触对胞体响应的影响。这一发现挑战了赫布理论在塑造神经环路中神经元选择特异性上的作用，并且表明了某一神经元

的选择特异性反映的是其突触前的一大群具有相似特性和不同权重的神经元的共激活作用。

电镜成像能够提供在突触水平上足够精细的组织结构与连接信息，而光成像技术则反映了这些精细结构的功能活动。因此光电联合技术的大力发展，为神经科学家探索一些神经活动的基础原则提供了更为直观的视角。

1.3 视觉神经元：小鼠视皮层中存在大量低特异性神经元

自从 Hubel 和 Wisel 发现了神经元的朝向选择性，长久以来的观点认为在哺乳动物的初级视觉皮层中，不同的神经元具有不同的选择特异性，这些神经元用以识别特定视野空间中具有特定特征的视觉信息。从初级视皮层到高级视皮层，这些神经元的选择特异性逐渐精细化，从简单的对特定朝向的纹理响应发放，到仅对特定的复杂特征响应发放。但是由于以往研究方法的局限性，人们通常仅对某一特定脑区中对视觉刺激有响应的神经元进行功能特性的检测。那么如果扩大神经元的采集范围至整个初级视觉皮层，或者全部视觉系统中的所有神经元，这一观点是否仍然成立呢？

针对这一疑问，Saskia 等人于 2020 年 1 月发表在 Nature Neuroscience 杂志上的工作[2]，对 Allen Brain Observatory 的开放数据进行了研究。这份名为‘the Allen Brain Observatory Visual Coding dataset’的数据集，包含了约 60，000 个视觉神经元的活动信息。这些神经元分别源自于 12 个转基因品系的 243 只成年清醒小鼠，涵盖了六个视觉区域、四个层内的神经元活动。这些神经元神经活动信号的获得，都遵循严格的采集流程，具有标准化的设备和实验方案，并且具有与刺激驱动响应无关的严格品控。在数据采集时提供的视觉刺激包括：光栅（移动的和静止的）、局部稀疏噪声、自然场景和自然场景视频。

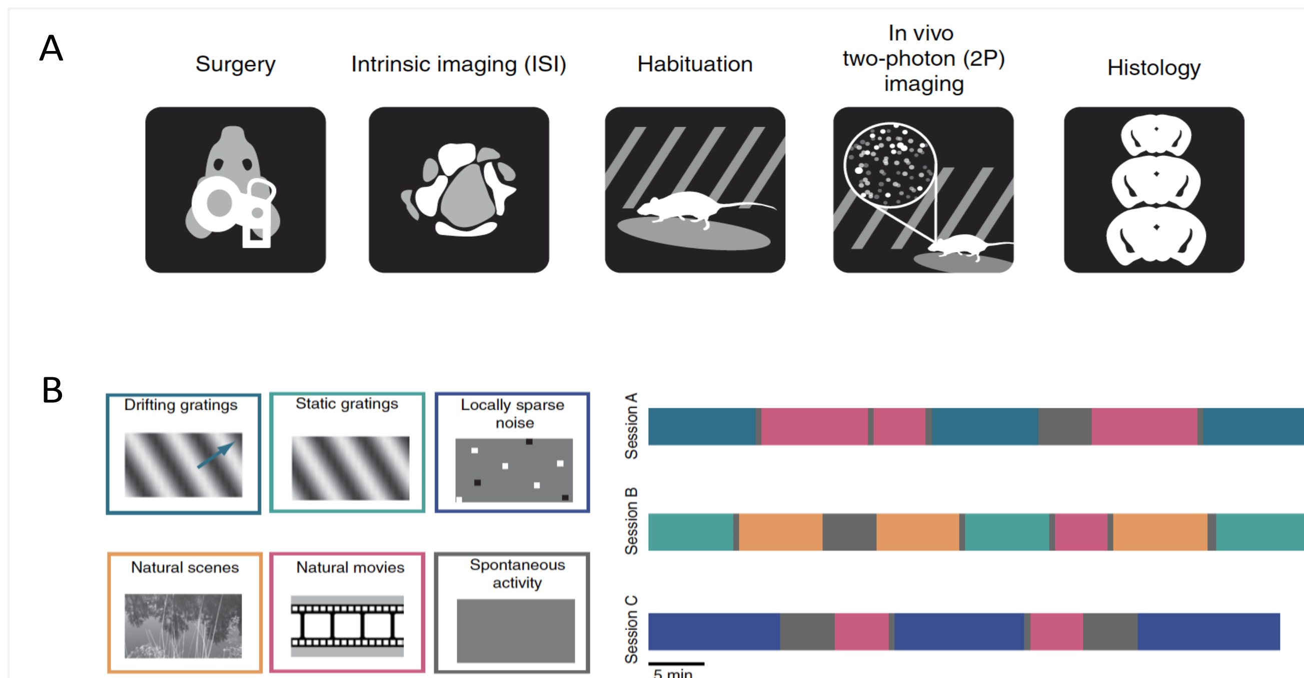


图 1.3 标准化数据采集流程及视觉刺激范式（图片源自[2]）

a)每一只小鼠进行大规模数据采集的工作流程示意图。b)用于检测小鼠视皮层神经元响应特性的视觉刺激的标准化实验设计。六组视觉刺激（图左）被分散在三段成像过程中（图右），这三段成像过程分不同天进行。

依据神经元对不同视觉刺激的响应活动，对这些神经元进行分类，并通过神经响应模型对神经元的功能响应分类进行验证。这一模型包含了视觉神经生理学建模中发现的大多数基本特征，以及小鼠的奔跑速度。结果发现，对于一类神经元而言，模型能够很好的预测其对人工视觉刺激和自然视觉刺激的响应。但是对于许多神经元而言，模型的预测并不佳。尤其是其中最大的一类神经元，它们似乎对任何刺激都没有稳定的响应，并且随着视觉皮层级别的升高，这一类神经元的数量在逐渐增多。这个结果表明，仍然有许多神经元并没有按照人们预期的模式响应发放，它们的响应模式多变且特异性低，而这一现象在高级视觉皮层中更为明显。这一实验展示了，获得大量的、无偏差的神经响应数据，对于准确理解神经系统的计算机制是十分重要的，也表明之前的视觉神经元-特定视野空间-特定特征的视觉信息模型的局限性。

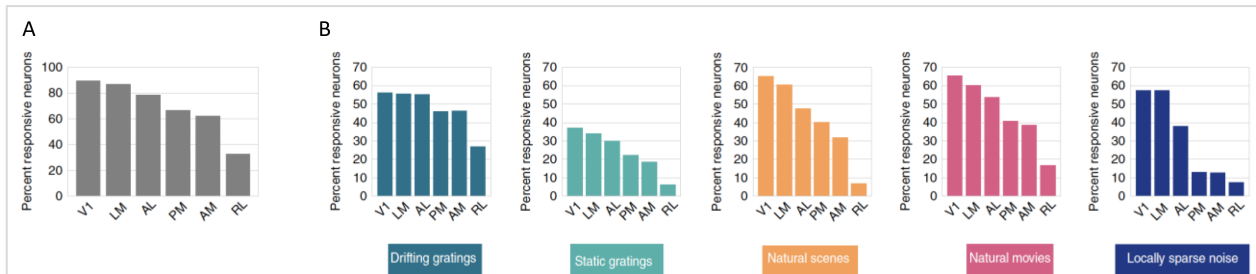


图 1.4 神经元响应视觉刺激的比例（图片源自[2]）

a)在不同视皮层中至少对一种视觉刺激响应的神经元的百分比。

b)不同视皮层中对特定类型视觉刺激响应的神经元的百分比。

1.4 视觉系统：小鼠视觉系统的功能层级

哺乳动物视觉系统的特点是具有层级结构。不仅是在解剖结构上，功能上也具有层级现象。源自于外界的视觉信息通过视网膜传递至高级视觉皮层，信息会经历层层加工、提取与传递。依据这一架构，人工神经网络的建立也获得了许多启发。AlexNet 是 Hinton 与其学生 Alex Krizhevsky 设计的一个卷积神经网络，并获得 2012 年 ImageNet 竞赛的冠军。这一网络的出现，将计算机视觉的研究带入了一个新的高潮。如今越来越多的研究发现，AlexNet 对图像的加工与人类的视觉系统有着很多相似之处。比如二者对物体特征的识别均从简单向复杂演化；而且在人脑视觉系统与 AlexNet 的不同层级中，对相同一组图片的表征相似性矩阵（RDM）也是十分相似的。但是由于技术上的挑战，同时记录哺乳动物多个脑区的神经活动较为困难，因此长久以来一直缺乏在细胞水平直接观察不同层级的功能以及跨层级的交互作用。2021 年 1 月，来自于艾伦研究所的 Joshua H. Siegle 等人，通过构建新型的神经电信号记录系统和神经像素电极阵列（Neuropixels），成功的观察到了小鼠视觉系统的功能层级活动。相关工作发表在 Nature 杂志上[3]。

在实验中，小鼠需要接受‘移动的光栅’‘全屏闪烁’和‘局部的 Gabor 光栅’等视觉刺激，并通过 Neuropixels 记录鼠脑中 6 个视觉相关皮层和两个丘脑视觉核团中约 100，000 个神经元的动作电位。在以往的研究中，人们并不确定鼠脑是否也如灵长类一样，存在视觉系统的功能层级。而本实验中的发现指出，鼠脑的 V1 神经元总是先于其它区域的神经元活动，AM 区域（antero-medial area）神经元总是晚于其它区域活动。这一现象暗示了小鼠的视觉系统也存在功能层级的结构。进一步的研究发现，随着层级的升高，各对应区域视觉反应的延迟也随之增加，感受野大小逐渐增大，其感受野的复杂度也逐渐增加。并且高级区域对反应的衰减时间也相对更长。而且实验还发现，这些神经活动与行为选择之间的相关性也沿着层级结构增加。应用新的实验技术，该研究为理解信息编码与信号在视皮层与丘脑视觉核团中的逐级传

递提供了基础。

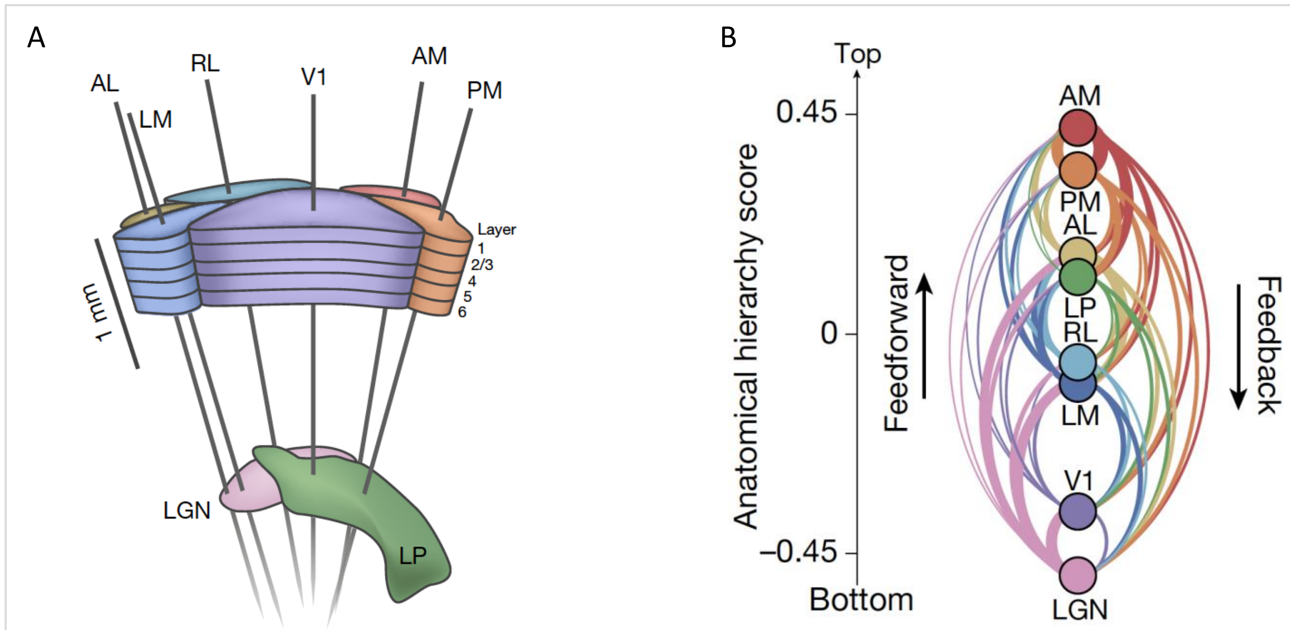


图 1.5 探针定位与解剖组织层级（图片源自[3]）

a)探针插入视觉皮层（V1, LM, AL, RL, AM, PM）和丘脑（LGN, LP）组织的轨迹示意图。

b)依据参考文献计算的八个感兴趣区域的解剖学层级评分。

随着电生理技术的不断优化，人们可以同时记录越来越多的神经元的神经活动。此时科学家对神经系统的探索，便可以不再局限于单一脑区的少量神经元。结合更多脑区间的相互作用信息，研究人员可以建立信息量更为丰富的系统模型。这不仅会拓宽人们对神经系统工作机制的理解，发现新的、或者以往无法直观观测的现象，有时还会对以往的理论提出新的观点。

1.5 全脑结构连接组：全连接组测序快速构建小鼠全脑连接组图谱

电镜成像技术，可以通过采集十分精细的脑组织结构信息并进行重构，来获得生物脑的全联接图谱。但是这样的工作将会消耗大量的时间与人力，因为成本昂贵，往往仅对少量样本进行，难以处理大量样本。而用传统的病毒示踪法绘制脑内长距离连接组图谱，单一样本中仅能够绘制少量的连接，要做到全脑连接组，便需要将多个样本的数据进行汇总。2020 年 7 月，黄龙文等人发表在 Cell 杂志上的工作[4]，提供了一种能够快速、高效、低价格实现全脑绘制长距离连接组的技术——BRICseq（Brain-wide individual-animal connectome sequencing）。

在这项新的全脑连接组测序技术中，研究者使用‘核酸条形码’作为示踪剂，来标记与追踪单个神经元的长距离投射。核酸分子具有无穷的多样性，使用核酸条形码标记神经元，不仅避免了传统方法中示踪剂颜色数目的局限性，还能够通过测序技术来提高长距离连接组的测绘通量。在实验中，研究者将携带有大量、多样且独特的核酸条码的病毒注入到小鼠的右侧大脑半球中。病毒感染神经元，并将扩增的核酸分子定位在神经元轴突的末端。经过扩增，神经元胞体中的核酸丰度要显著高于其在轴突中的丰度。因此通过测序技术，便能够获得哪些脑区中包含同一个神经元的胞体及轴突（即，可以检测到相同核酸条码），以及这一神经元是从哪一个脑区投射至其它脑区的。将充分进行病毒表达后的鼠脑分割成若干小块，然后通过对每一个小块中的核酸条码进行测序分析，便可以获得小鼠的全脑长距离连接组。通过将 BRICseq 技术应用于小鼠的新皮层，研究人员发现少量基因的空间表达模式，能够预测皮层间的长距离连接。

而皮层间解剖学上的连接强度，则与皮层间的功能连接强度十分相关。这一结果提示了大脑的基因表达、连接模式和神经活动间可能存在简单而直接的关联。另外，研究者还利用 BRICseq 技术重现了 BTBR 品系小鼠的脑连接缺陷。这意味着该项技术将在研究检测一些神经连接缺陷的疾病动物模型中，发挥重要的作用。BRICseq 技术的出现，将使得全脑尺度的基因表达、结构连接、神经活动和行为学数据能够整合在一起，从而更为全面的探究脑的工作机制。

生物脑内大量神经元间的长距离与短距离连接，为生物的脑网络提供了“硬件”基础。虽然以往的研究已发现，神经系统的运行受到多方面因素的调节，单纯的获得硬件连接的信息，还远不足以充分解析生物智能的工作机制。但正如上文中对单一神经元精细结构的获得，有助于人们更好的理解神经活动的基础原则一样；开发快速、方便的全连接组图谱绘制技术，能够在全脑尺度上高效、大规模的获得神经元间，以及脑区间的组织连接关系，必将为人们理解神经系统的工作机制提供有力的支撑。

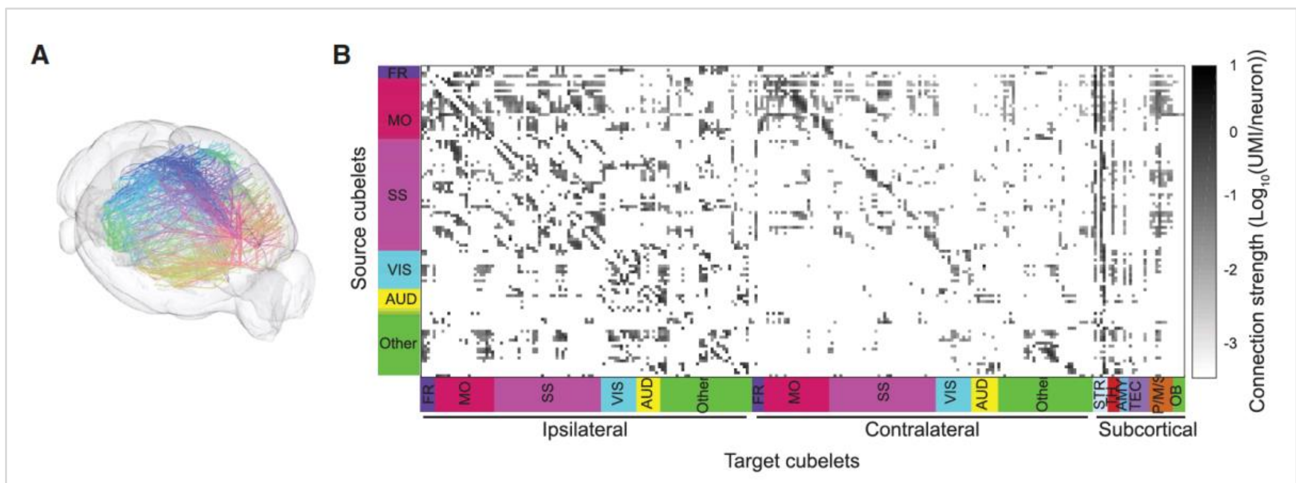


图 1.6 BRICseq 技术绘制的全脑皮层连接组及各区域连接强度分析（图片源自[4]）
a)BL6-1 小鼠全脑皮层中分割出的组织小块与组织小块间的连接示意图。b)鼠脑分割出的各小块组织间的连接强度分析。纵坐标代表源组织，纵坐标代表目标组织。图中已按照各组织所属的脑区进行了分类。

1.6 全脑功能连接组：高通量全脑成像辅助预测斑马鱼决策行为

由于技术上的挑战，以往对生物脑的研究多集中在局部脑区。随着光学成像技术的兴起，结合斑马鱼幼鱼这样一种神经网络规模适中，且脑组织高度透明的模式生物，使得高通量的全脑成像成为可能。并且随着光片、光场显微成像技术的逐步发展，人们可以在更高的时间分辨率上获得生物全脑的神经活动信息。这就使得人们可以实时检测完整的生物脑在进行信息加工时的活动，这将为理解神经系统的工作机制提供更为丰富的信息。对生物智能过程的尽量如实的还原，也将有助于对生物智能进行仿真，为开发功能更为优化的类脑神经网络提供助力。

2020 年 1 月 Qian Lin 等人在 Cell 杂志上发表的工作[5]，通过在全脑范围内对目标引导的决策行为进行研究，揭示了斑马鱼的小脑在认知功能上的作用。在外界刺激的作用下，生物会形成趋利避害的学习记忆，然后通过目标引导的行为对不同的外界环境作出相应的响应。Qian Lin 等人通过建立操作条件化范式，来研究斑马鱼幼鱼运动选择行为的神经基础。在实验中，实验者会在斑马鱼幼鱼的头部放置一个热源，通过加热向幼鱼施加一种厌恶刺激。当幼鱼的尾巴向目标方向（向左或向右）摆动时，热源的加热便会停止，否则（其它尾动情况）则会持续加热。经过少量试次的训练，幼鱼能够学会这一行为。这一范式涉及到了感觉输入、信

息整合、学习记忆、决策和目标引导的运动等多种神经认知过程。

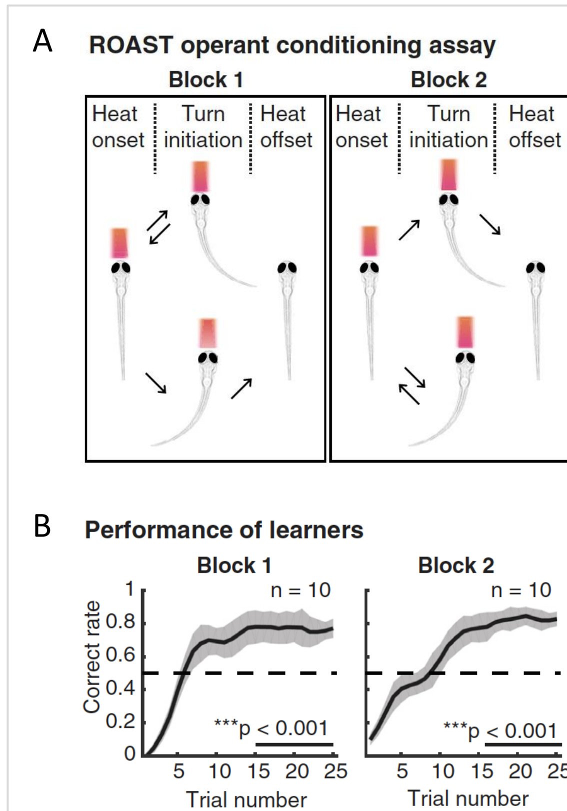


图 1.7 斑马鱼的操作条件化范式
（图片源自[5]）

a)斑马鱼幼鱼的操作条件化范式。幼鱼的头部施加一个红外线激光热源，如果幼鱼在刺激呈现时向目标方向摆尾，则会诱发热源关闭，否则热刺激会持续呈现。在 block 2 中，目标方向会切换至另一侧。

b)无论是 block 1 还是 block 2，幼鱼均能够快速学会向目标方向摆尾，以减轻厌恶刺激。其准确率在 15 个试次左右趋于稳定。

在执行任务的过程中，实验者通过自制的光场显微镜，在幼鱼的全脑范围内实时采集神经活动。经过对多个脑区中约 5000 个最为活跃的神经元进行追踪，发现参与这一行为任务的神经元分布在整个前脑、中脑和后脑前部。由此可见这一行为的执行，需要整合多个脑区的信息。依据 t-SNE 非线性降维分析，这些神经元可被划分成 6 个具有不同功能特征和解剖区域定位的神经元类群。每一个类群都参与到这一行为任务的不同阶段。比如类群 1 与热刺激的出现有关，类群 6 与热刺激的消失有关。经过分析这些神经元的响应模式，实验者能够依据神经活动在单个试次中运动发起前的 10s 以上，预测幼鱼即将发起的运动方向。他们发现，在一个个试次中，幼鱼的“脑状态”（即所有神经元在特定时间的活动状态）呈现周期性的动力学变化。在执行运动前，不同运动方向的决策会呈现出显著不同的脑状态。依据这些神经活动上的差异，能够成功的预测幼鱼接下来的尾动行为，无论在这一个试次中幼鱼决策时间的长短。然后在新的试次中，脑状态变化又会进入新的循环。这些神经活动的动力学变化产生于分布式神经网络，这一网络的功能连接会在试次间发生变化。人们发现在执行行为任务正确的试次中，在运动发起前，幼鱼同侧小脑（ipsi CB，ipsilateral cerebellum）与缰核（Hb，habenula）间的神经活动的相关性，要显著高于错误的试次。对于形成正确决策的运动计划，脑区内与脑区间的信息整合与调节是十分必要的。经过进一步的分析认为，幼鱼的小脑和 ARTR 区域（anterior rhombencephalic turning region）中的信号，对于尾动的方向决策具有最强的预测性。在进行运动计划过程中，小脑同侧和对侧半球的神经活动差异构成了方向决策的神经基础。神经响应更强的一侧，决定了方向选择的输出。而双侧小脑的群体神经活动编码了每一个试次中神经系统作出决策的时间。

在以往的研究中，人们普遍认为小脑主要与运动相关。而这项研究则发现了小脑的认知作用，及其在运动计划中的重要性。在实验中，人们还发现小脑中的颗粒细胞的神经活动对于运

更多资料，欢迎访问 www.wenku120.com

动的决策预测是十分重要的。在以往的研究中，人们曾发现颗粒细胞也参与到学习记忆之中。可见这一实验使人们对小脑颗粒细胞的功能有了新的了解。此外这一实验结果还反映了，目标引导的行为需要多个脑区的共同参与，因此在日后的研究中，更多的获得跨脑区的神经信号，将有助于我们更全面的探索感觉运动整合的机制。

斑马鱼幼鱼的优势在于能够进行高通量的在体光学成像。新型光学成像技术的不断优化，不仅在成像视野的水平上获得了提高，在成像的空间分辨率与时间分辨率的纬度上，也获得了较大的突破。这使得研究者不仅能够直接观测到单个神经元的活动，还能够更准确的分析神经活动的时间关系。这使得人们向更高还原度的检测神经系统的信息加工过程迈进了一步。

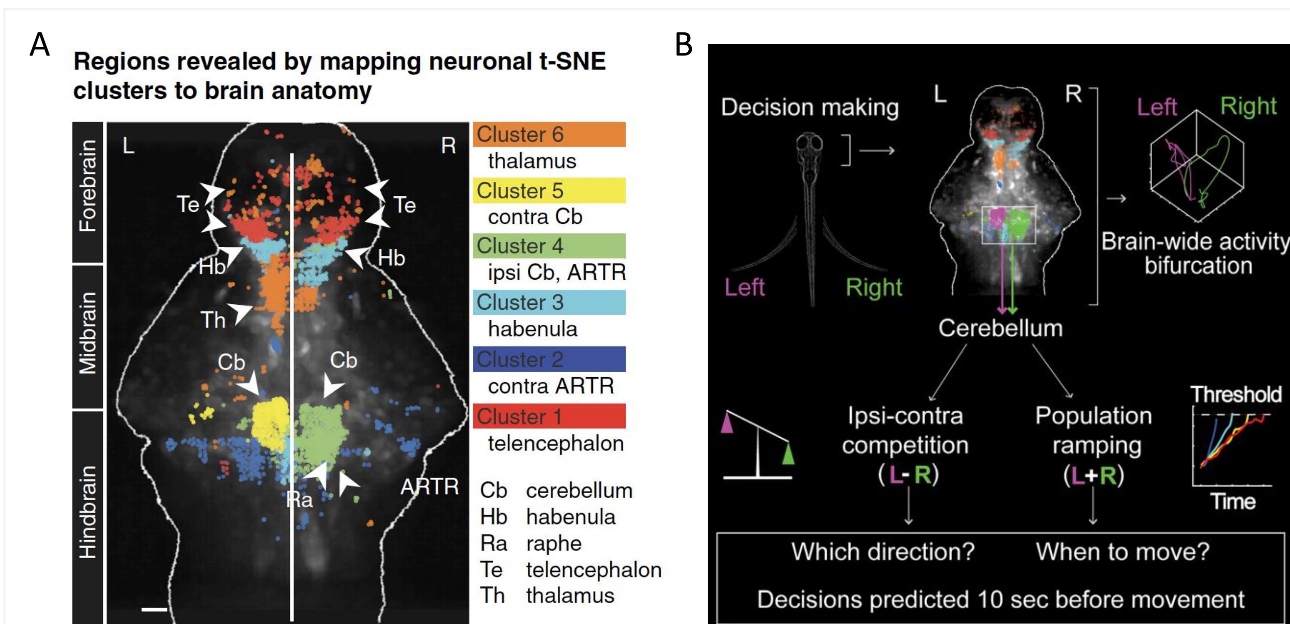


图 1.8 参与行为加工的神经元位于的脑区及工作机制示意图（图片源自[5]）

a)t-SNE 非线性降维分析划分的神经元类群，及其在斑马鱼幼鱼脑中的定位。b)斑马鱼小脑神经元的活动信号可以在运动发生前 10s 预测决策方向。同侧与对侧小脑半球神经活动的竞争决定鱼尾运动方向，两个半球的神经元群体的神经活动共同编码决策时间。

1.7 认知理论: 新证据引发修订传统认知理论的思考

随着脑成像、神经信号采集、神经调控等相关技术手段的更新与发展，人们对大脑的工作机制有了更深入的了解。与此同时，研究者们也发现了一些难以采用传统认知理论合理解释的新证据，例如 2020 年发表在《Trends in Cognitive Sciences》的两篇文章，分别就视觉通路[7]和人类海马神经元记忆编码[8]提出了新的观点和建议。

传统模型认为，灵长类动物的视觉系统可以分为两个通路：以颞叶为代表的腹侧通路（what 通路，负责辨认客体），和以顶叶为代表的背侧通路（where 通路或 how 通路，负责客体定位或为抓握、操作做准备）。然而，David Pitcher 等在《Trends in Cognitive Sciences》发表观点称，越来越多的证据提示人类和非人灵长类动物大脑外侧表面存在第三条视觉通路[7]。在解剖结构上，这条通路起始于早期视皮层，经由运动选择区（V5 / MT），投射到颞上沟（STS）。与背侧和腹侧通路隔离（图 1.9）。在功能上，第三条视觉通路也与其他两条通路不同，它选择性地对运动中的面孔和身体做出反应，可根据动态的社交线索计算一系列较高级的社会认知功能（例如表情、注视、视听融合、意图和情绪等）。基于这些证据，David 等认为，第三条视觉

通路是计算面部信息以支持动态且不断变化的社交互动过程的关键，有必要在传统的背侧和腹侧视觉通路模型的基础上添加第三条通路的定义。

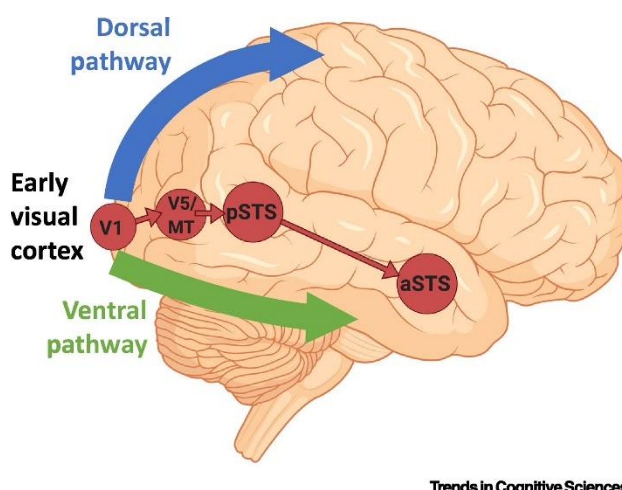


图 1.9 第三通路的皮质连接（红色）（图片源自[7]）

另一项颠覆传统认知的研究是关于人类海马记忆编码原理的探讨。在以往的研究中，来自理论、模型以及大鼠和猴子的单神经元记录实验等证据表明，模式分离是大鼠和猴子海马记忆编码的基本原理，而人类差异性 fMRI 激活的证据也反映了模式分离的过程。但是，莱斯特大学研究者 Rodrigo Quian Quiroga 基于人类单神经元信号采集的证据在《Trends in Cognitive Sciences》发表观点称，人类海马中的记忆编码可能不存在模式分离[8]，这一结论对解释人类独具的认知能力（比如概括能力、创造性和抽象思维能力等）具有深远的影响。在该项研究中，Rodrigo 等人对内侧颞叶（MTL）植有电极的药物难治型癫痫患者进行了单神经元电信号采集，结果显示某些细胞只对特定的概念有反应（称为“概念细胞”）。例如，图 1.10A 展示的神经元仅有选择地对演员成龙的面孔及其名字响应，对其他面孔和物体则没有反应。此外，人类 MTL 神经元也可以对有关联的概念响应。如图 1.10B 中的神经元，对演员尼尔森的面孔图片和飞机的两张图片都做出了响应。这两张在视觉特征上几乎没有相似性的图片，都与尼尔森主演的电影《飞机》有关，属于相关联的概念。进一步分析发现，对同一概念的不同形式或两个关联概念有反应的 MTL 神经元，在响应强度和潜伏期上都是难以区分的。Rodrigo 认为这体现

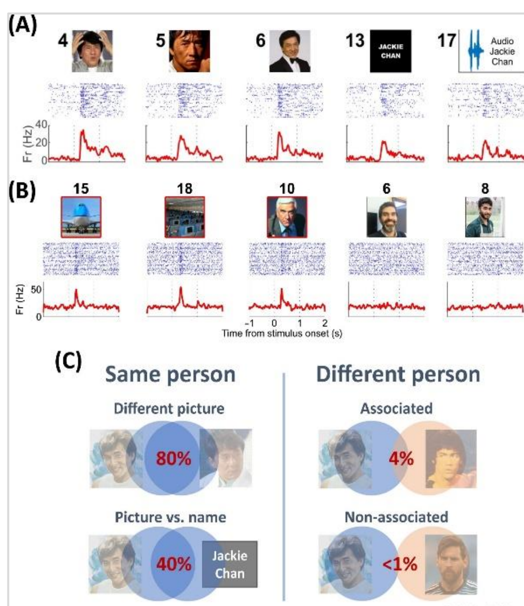


图 1.10 人类 MTL 神经元对概念和关联的编码（图片源自[8]）

(A)海马区某神经元对成龙的三张面孔图片、成龙的姓名音频以及成龙的姓名文本图片的响应。
(B)海马区另一个对飞机图片及尼尔森的面孔图片的响应，前三个图片与尼尔森主演的电影《飞机》有关，该神经元对后两张无关图片无响应。
(C)MTL 神经元对成龙的不同图片、成龙的面孔图片与姓名文本图片、不同但相关的两个人（成龙和李小龙）的图片，以及不同也不相关的两个人（成龙和梅西）的图片的经验发放概率。

了一种类似二进制的编码机制，即 MTL 的神经元大多以同样的发放模式响应相同身份的面孔图片（因为它们在记忆功能上所传达的含义相同）；针对很相关的两种身份的面孔则保持不同的响应形式，但概念神经元可以将它们联系在一起（图 1.10C）。

在 Rodrigo 的研究中，另一个有趣的结论是可以通过让被试学习概念间的新关联来调节 MTL 神经元的响应。例如，图 1.11A 展示了某个神经元最初对演员布洛林的面孔（P）响应，而对单独的埃菲尔铁塔图片（NP）没有响应。而在被试观看了带有布洛林的艾菲尔铁塔图片后，得知了这两种刺激之间是相互关联的，之后该神经元也开始对没有布洛林的艾菲尔铁塔图片做出响应（图 1.11B）。

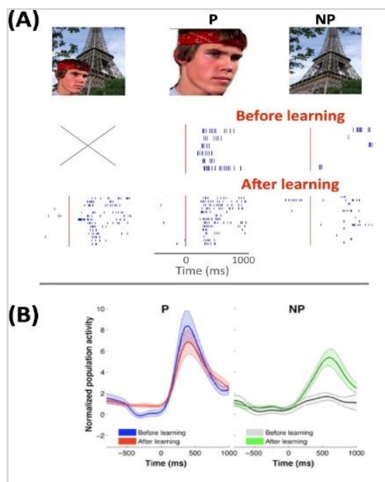


图 1.11 人类 MTL 神经元对新关联的编码（图片源自[8]）

（A）海马区某神经元对布洛林的面孔图片（P）有响应，但对埃菲尔铁塔图片(NP)没有响应，在得知二者有关联后，该神经元对埃菲尔铁塔也有响应。
（B）成功编码 P 与 NP 之间的关联的神经元在学习这种关联前后的归一化平均响应情况。

1.8 总结与展望

在过去的一年里，随着新型实验技术的发展与应用，我们对认知功能的神经基础从微观到宏观都有了突破性的认识。如 1.1 所介绍，光学成像与电镜成像技术相结合，揭示了“强突触”和“弱突触”在视觉选择特异性的树突计算过程中同样重要。又如 1.7 中，通过回顾积累多年的人类在体单神经元结果，研究中提出海马中可能存在可以进行学习的，不拘泥于感觉输入模态的“概念细胞”。然而这些突破，对于我们庞杂的认知功能来说，只属冰山一角，未来尚有大量的“谜团”亟待探索。对于未来的发展趋势，我们有如下展望：

（1）高还原脑仿真：随着人们对大脑不同层次、类型的信息的获得，将参与形成以及调节神经系统功能的各项因素进行整合，并进行建模将成为未来研究生物智能的一个趋势。因此结合各领域的发展，未来系统神经科学的一个发展方向，将是充分利用能够获得的各层次、各方面的神经信息，通过将全脑范围内的连接图谱信息、精细组织结构信息，以及不同层次的神经活动信息，和同时采集的神经递质的动态变化信息充分的结合在一起，并进行建模，以获得充分还原神经系统真实活动的模型，并从中解析各项神经机制的运行原理。这些信息的获得，不仅能够帮助我们更好的理解生物智能的涌现机制，还能够验证认知神经科学原理，为人工智能的研究提供启发与参考。同时也能为理解各项脑疾病的发病机理，以及开发新型预防、诊断、和治疗手段提供帮助。

（2）向高级灵长类转化：在体清醒灵长类动物实现同时对神经递质、突触、神经元、功能柱、甚至全脑观测技术，解决更多的认知功能神经基础问题。以往在低等模式生物中的研究，对神经科学的发展起到了十分重要的推进作用。但是对于研究一些高级的认知行为，就需要以更为高级的模式生物作为样本。因此认知与神经科学的一个可能的发展方向，是在将各层次的神经活动采集技术向高级灵长类转化的同时，验证一些在低等模式生物中发现的机制与原理在

高等生物中是否保守；并探索更高级的认知行为的工作机制与形成原理。这些机制的发现，将为提供更接近人脑的模型、以及启发具有更高级认知功能的人工智能提供帮助。

（3）新认知理论模型建立与修订：随着认知神经科学领域新发现的不断出现，一项可能的认知神经科学发展方向是结合不断涌现的新证据，建立、提出新的认知模型，或补充新的认知理论。如，开发第三条视觉通路对应的人工神经网络模型。计算科学与神经科学的交叉，同时为人工智能和认知神经科学带来长足发展，二者在相互借鉴和相互促进中得以螺旋提升。基于认知科学理论，深度神经网络技术已经能在计算机内模拟传统的视觉通路，帮助研究人员更深入地理解大脑的工作，了解更多相关疾病背后的机理。未来，通过进一步明确第三条视觉通路，并探究其相应的人工神经网络模型，可能会带来更高层级的人工智能（涉及社会认知/社会交互）应用。这类应用反过来又能帮助我们明确新的认知神经科学问题和研究方向，以更加深入地理解大脑的工作机制。

参考文献

[1] Scholl, Benjamin, et al. "Cortical response selectivity derives from strength in numbers of synapses." *Nature* 590.7844 (2021): 111-114.

[2] de Vries, Saskia EJ, et al. "A large-scale standardized physiological survey reveals functional organization of the mouse visual cortex." *Nature neuroscience* 23.1 (2020): 138-151.

[3] Siegle, Joshua H., et al. "Survey of spiking in the mouse visual system reveals functional hierarchy." *Nature* (2021): 1-7.

[4] Huang, Longwen, et al. "BRICseq bridges brain-wide interregional connectivity to neural activity and gene expression in single animals." *Cell* 182.1 (2020): 177-188.

[5] Lin, Qian, et al. "Cerebellar neurodynamics predict decision timing and outcome on the single-trial level." *Cell* 180.3 (2020): 536-551.

[6] Sabandal, John Martin, Jacob A. Berry, and Ronald L. Davis. "Dopamine-based mechanism for transient forgetting." *Nature* (2021): 1-5.

[7] Pitcher, David, and Leslie G. Ungerleider. "Evidence for a third visual pathway specialized for social perception." *Trends in Cognitive Sciences* (2020).

[8] Quiroga, Rodrigo Quian. "No Pattern Separation in the Human Hippocampus." *Trends in Cognitive Sciences* (2020).

第二部分 新技术在认知与神经科学领域的应用

脑科学很大程度上是一门由实验技术驱动的科学。从最早的高尔基染色技术，到后来的单细胞电生理记录、功能核磁共振、双光子成像、光遗传技术等等，每一次技术进步都有力地推动了人们对于大脑的结构与功能的认识。因此，了解脑科学新技术的进展，是把握这一学科发展方向最为有效的方法。本部分围绕过去一两年以来，认知与神经科学领域新的技术突破及其应用，特别是聚焦成像与神经调控技术的重要进展做简要介绍。

脑是一个以层级化架构组织起来的复杂系统，神经科学需要通过先进的技术手段从全脑到单细胞乃至单突触的各个空间尺度，以及从分钟到毫秒的各个时间尺度观测大脑活动以揭示其工作原理。本部分的前一半内容将把重点放在跨尺度神经系统活动观测的核心技术之一——成像技术的开发与运用上，主要从新的成像硬件方法与新型荧光探针两个方面进行介绍。

在观测技术之外，对于神经系统的活动进行精确受控的扰动，以实现特定功能的技术，即神经调控技术，也是揭示脑工作原理的重要技术手段。一方面神经调控技术可以阐明脑活动与脑功能之间的因果关系（观测技术本身只能发现相关关系），另一方面由于它对脑功能的调节，也是脑科学技术中最有望对于脑疾病患者，甚至普通人带来帮助的技术。过去一年中，在视觉认知功能增强领域，非侵入式和侵入式神经调控技术在猕猴和人类的实验中取得了突破性进展。本部分的后一半将介绍两项成果，分别通过初、高级视皮层植入高密度电极阵列和实施新型颅内动态电刺激技术，实现视觉知觉的重建。这种颅内动态电刺激技术在与人工智能领域的新发展相结合后，有望实现更为复杂的视觉认知功能增强。

2.1 成像技术在神经科学领域的应用

光学显微成像设备是最主要的脑成像设备之一。近年来，在更深层组织成像、更大视场成像以及更高时空精度成像等方向上，光学显微技术都取得了很大进展。本文涉及的光学显微技术包括多光子显微技术、宽场显微技术以及光片显微技术等。

具体而言，多光子显微镜应用广泛，具有高对比度、低光毒性和良好的深层成像能力，但成像速率受限制。本部分将在 2.1.1 与 2.1.3 中讨论多光子显微镜针对应用场景，成像深度，成像速率的新方法与技术改进。传统宽场显微镜视场大，成像速率高，但是体成像与多层成像能力太差，2.1.2 将介绍宽场显微技术在体成像上的进步。2.1.3 将介绍新型光片显微技术，能够用于本领域的高速大视场体成像。2.1.4 将介绍自由活动的生物成像，这是研究神经活动与行为关系时非常关注的技术问题。

2.1.1 多光子深层成像技术

相比传统的单光子共聚焦显微镜和常规明场显微镜而言，多光子激光扫描显微镜在不透明的生物组织中的成像深度更深，对样品伤害更小，因此最适合观察模式动物大脑在发育、疾病等过程中的细胞结构与神经活动变化。而新的多光子技术正在改变传统的实验方式，下面将介绍两种新型双光子/多光子设备。

双光子成像与虚拟现实记录结合：2020 年，冲绳科学技术大学院大学（OIST）的 Kuhn 等人将头部固定的小鼠双光子成像技术与虚拟现实系统和行为测量系统结合，从而解决了长时间检测记录小鼠大脑神经活动的问题[1]。研究人员将每只小鼠放在显微镜下的轨迹球上，固定

头部，并在其周围投影一个虚拟现实迷宫。当老鼠在虚拟迷宫中奔跑时，就可以记录下它的神经活动。Kuhn 与萨特仪器公司（Sutter Instrument）历时五年共同开发了一个成像平台，其核心是可移动物镜显微镜（Movable Objective Microscope, MOM）（图 2.1），通过移动物镜实现空间扫描。同时，该平台能够用摄像机记录至多 3 小时动物行为，并使用公司自有的 MScan 软件来控制实验和整合数据。



图 2.1 可移动物镜显微镜（MOM®）是一种双光子或三光子显微镜，当与适当的激光器组合使用时，能够对活体标本进行深层成像。 Sutter MOM 是第一个能够 3 维移动和旋转的物镜，物镜沿 X，Y 和 Z 轴平移并绕 X 轴旋转。并且允许物镜通过旋转离开标准垂直位置自由定位至 0 到 180 度之间任意角度（如右图所示）。这种位置自由度允许对非水平表面和物体进行成像，无需移动台或其他方式干扰样品。即可在保持样品稳定的同时记录大范围，不同角度与深度的组织图像。（图片源自[2]）

运用这套方法，Kuhn 团队成功的记录了时长两小时的老鼠的复杂行为与脑部神经活动（图 2.2）。他们同时还测量了老鼠警觉度、舔胡须频率等行为学参数。Kuhn 团队将两个多小时的行为学数据记录与 MOM 采集的神经活动进行了匹配，该过程中小鼠逐渐疲劳并入睡，因此能够观察到小鼠不同阶段的不同的行为[1]。

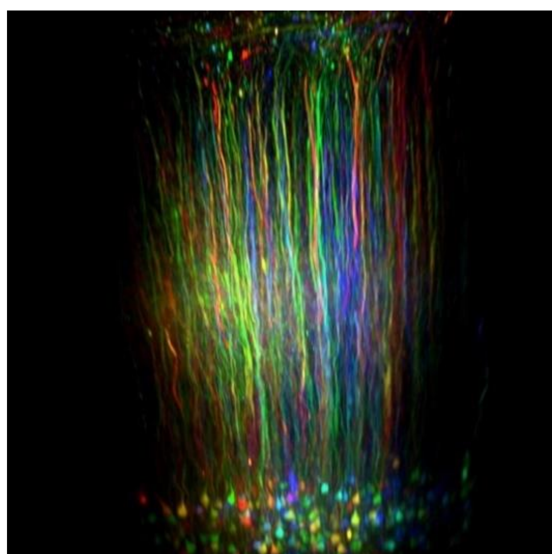


图 2.2 使用可移动物镜显微镜（MOM）拍摄的运动小鼠的视觉皮层神经活动。
（图片源自[3]）

三光子显微成像：2020 年，多光子显微技术在更深层成像的方向也取得了进展。研究人员开始用三个波长更长、能量更低的光子激发更深层的荧光团。双光子显微镜使用通常可以穿透 0.5-1.0 毫米的脑组织。康奈尔大学的物理学家 Chris Xu 和他的同事使用了波长为 1700 纳米的三个光子作为激发光，可以穿透 1.5 毫米的脑组织而不发生散射。最近 Xu 与合作者使用三光子显微镜给一条活体成年斑马鱼的大脑进行成像（图 2.3）。斑马鱼是常见的研究大脑发育的模式生物，它们的头骨在幼年时是半透明的。但随着年龄的增长，鱼体变得不透明，Xu 等人实现了在一定条件下，使用三光子显微镜透过成年斑马鱼头骨看到大脑内深度约 0.5 毫米处的组织图像[4]。这项技术有助于更好地比较成年斑马鱼和幼年斑马鱼大脑的区别，为斑马鱼大脑发育提供新的研究手段和底层数据。三光子显微镜技术也将用于研究大型动物的大脑。

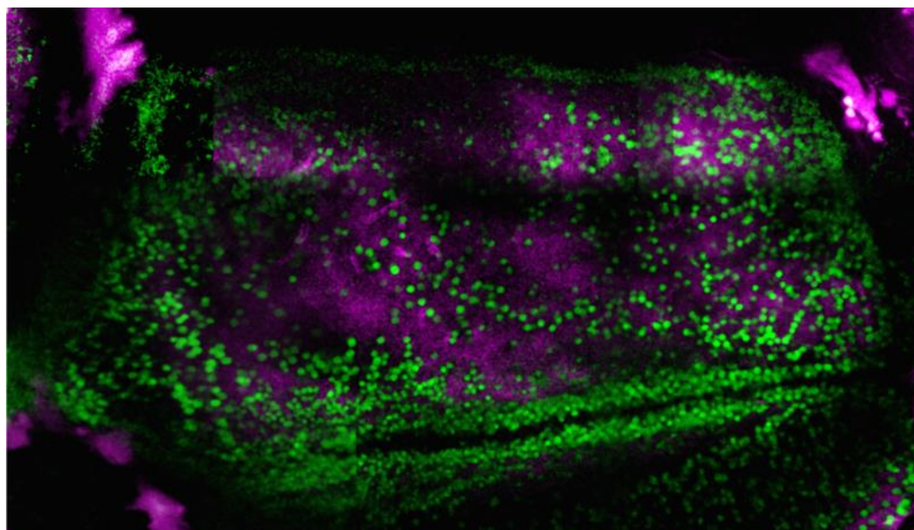


图 2.3 三光子显微镜能够透过成年斑马鱼的头骨观察神经活动。（图片源自[3]）

2.1.2 超大视场成像

宽场荧光显微镜（wide-field fluorescent microscopy, WFFM）由于其特有的高速记录能力和成本优势，被经常用于在大视场下对生物动力学现象进行光学成像。但是神经生物学中体成像或者多层成像要求显微镜具备光学层切能力。与双光子显微镜等点扫描显微技术不同，传统的 WFFM 一般不具备光学层切能力（或较差）。改进的 WFFM 如结构光照明显微镜（structured illumination microscopy, SIM）实现了宽场光学层切（wide-field optical sectioning），从而使得宽场显微镜对生物样本进行多平面的三维体成像技术上成为可能。但是，传统的轴向扫描是通过机械移动物镜或生物样本来实现的，这导致轴向最大速度将受到物体惯性的影响，不可能实现高速轴向扫描，从而限制了宽场显微镜的高速三维体成像能力。

对于具备光学切片能力的宽场显微镜，可以采用 ETL 实现无需机械扫描的快速多平面成像与体成像。该技术使了一种基于电可调焦透镜（Electrically tunable lenses, ETL）的新的轴向扫描方法，实现快速的轴向扫描。该技术首先在结构光照显微镜（SIM）上实现[5]，但样品动态变化会导致 SIM 重构后的图像容易出现运动伪影。为了克服这些问题，2019 年，清华大学的戴琼海团队开发了基于宽场荧光显微镜的多层高速活体内成像技术[6]。这种新型的多平面宽场荧光显微镜具备光学层切和体成像能力，以及 1.56 mm×1.83 mm 的超大视场（field-of-view, FOV），达到了 6.4 μm 轴向精度及 1.49 μm 的横向分辨率，适合对神经细胞胞体成像。

使用这套宽场成像系统，戴琼海团队成功的对活体内生物动力学现象进行了多平面成像。

他们以 5 赫兹频率对清醒小鼠的 FITC 染色的脑血管进行双平面成像（图 2.4）；对活体斑马鱼全脑钙信号在三个不同深度的平面进行成像（图 2.5）。成像结果表明，他们开发的成像方法相对普通宽场成像更灵敏，背景荧光更弱，有可能应用于观测三维大视场下的神经活动。

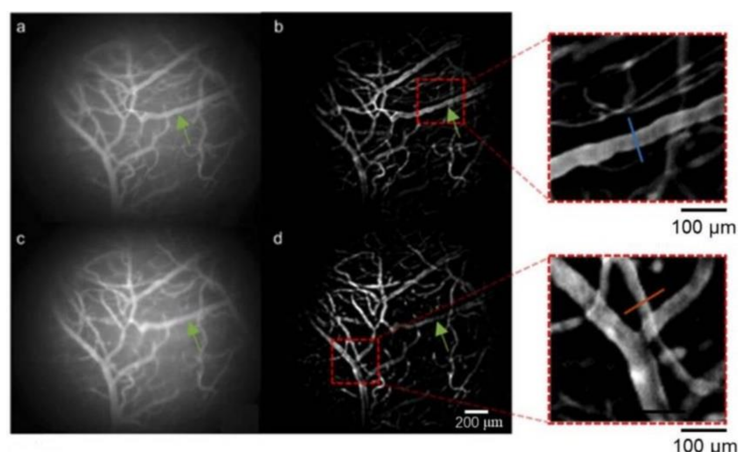


图 2.4 小鼠大脑皮层血管扩张的双平面活体成像。其中左列为均匀照明宽场成像，右列为新型多平面宽场显微镜成像，子图为细节放大图。（图片源自[6]）

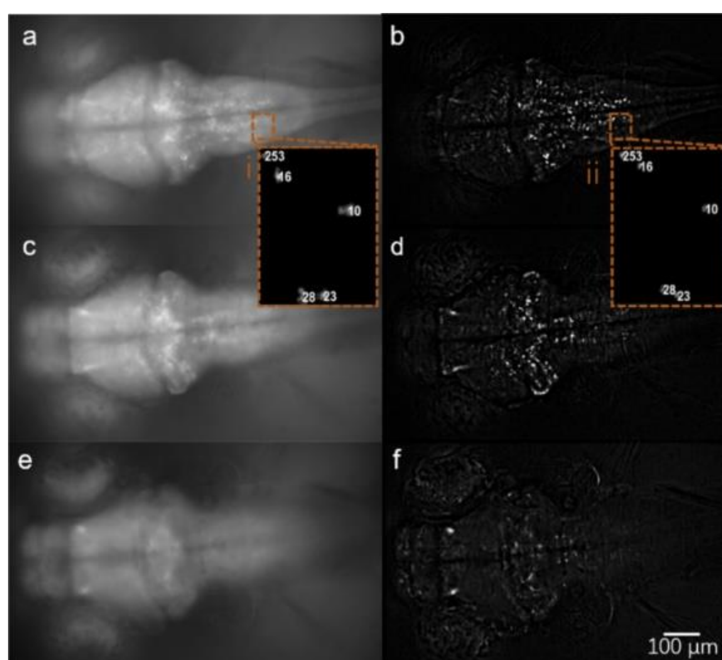


图 2.5 斑马鱼全脑三平面活体成像。其中左侧为均匀照明宽场成像，右列为新型多平面宽场显微镜成像，子图为细节放大图。（图片源自[6]）

2.1.3 超高速成像

常规的点扫描双光子显微镜的成像速率受到振镜的限制，只能达到每秒数十帧的水平。要在大脑深处以高时空分辨率捕捉膜电势变化等快时间尺度事件，人们需要高速双光子显微成像技术。另一方面，光片显微镜的高速成像能力也在不断被完善。

全光无源激光扫描高速成像：2020 年，加州大学伯克利分校的 Na Ji 团队开发了超快双光子荧光显微技术，达到了每秒 3000 帧的成像速率和亚微米级别的空间分辨率，可以在活体

内对神经元活动进行成像[7]。Ji 团队利用基于自由步角-chirp-增强延迟的全光无源激光扫描仪（all-optical passive laser scanner based on free-pace angular-chirp-enhanced delay，FACED），开发了具备 1000 赫兹和 3000 赫兹点扫描频率的双光子显微镜，并提高了荧光信号信噪比。他们将这项技术应用于对胞内钙离子活动、谷氨酸释放和膜电位的超快观测。该超高速双光子显微技术也可用于亚细胞尺度成像。使用 FACED-双光子荧光显微镜对荧光标记的钙活动进行 1 千赫兹速率成像，神经轴突就分辨，并且该技术能够检测到离体培养神经元和小鼠脑切片中的瞬时钙活动，甚至能够记录轴突中以 25μm/s 速率传播的自发钙信号（图 2.6）。值得一提的是，添加 FACED 模块后，基于标准振镜的双光子荧光显微镜便可轻松转换为千赫兹成像系统，这意味着该技术的广泛适应性。

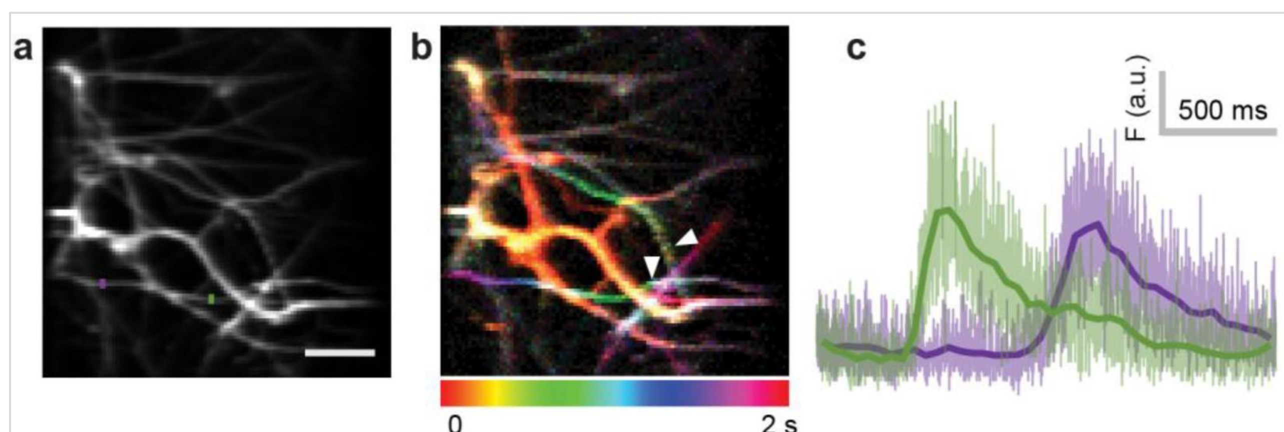


图 2.6 表达 GCaMP6s 的离体培养神经元的神经突中自发钙信号成像。
a. 2000 帧图像的平均强度投影。b. 颜色表示的钙信号传播图。白色箭头指钙信号最开始增强的位置。
c. 在 a 中的绿色与紫色两点处的钙信号随时间的变化，计算表明钙信号的传播速度约为 25μm/s。
图中比例尺为 10μm。（图片源自[7]）

基于轴向扩展的贝塞尔光扫描双光子显微镜：同年，Ji 团队又开发了一种基于轴向扩展的贝塞尔光双光子激光扫描显微镜（two-photon laser scanning microscopy with an axially extended Bessel focus，Bessel TPLSM），实现了大脑中三维空间血管循环系统的结构高速成像[8]。Bessel TPLSM 可以在大视场下直接分析血管的扩张和收缩动作，并以 99 体每秒（volumes per second，VPS，指每秒对样本进行体成像的频率）的速率监测血流活动。Bessel TPLSM 技术将对神经血管研究产生广泛的影响。

在体单镜头扫描光片显微镜：光片显微镜是另一大类体成像技术，具有高扫描速率和低光损伤，然而传统的多物镜光片显微镜在全脑成像领域运用很有限。2015 年，Bouchard 等人在开发了扫描共聚焦单镜头光片显微镜（swept confocally aligned planar excitation，SCAPE）[9]，突破了传统光片显微镜的局限性。该技术将振镜与倾斜平面照明相结合，使激发光片可以横向扫过样品，同时保持相机焦点始终在运动平面上，其过程如图 2.7a 所示。与点扫描方法相比，SCAPE 显微镜还具有两个重要优势。第一，SCAPE 的物镜在 3D 成像过程中保持完全静止，从而消除了外界干扰或样品反应的可能，并能够集成其他用于样品观察或光遗传学操作的光路；第二，它能够通过其独特光路以很高的速度捕获立体图像。

高速光片系统在神经科学中具有巨大的应用潜力。2018 年，Hillman 等人利用这种技术以 10VPS 的时间分辨率捕捉了清醒的正常行为小鼠大脑树突的钙动力学（图 2.7b），这相当于以 156 兆像素每秒的速率进行双光子显微镜成像[10]。并且 SCAPE 显微镜还可以结合使用双光子技术可以进一步提高成像深度。

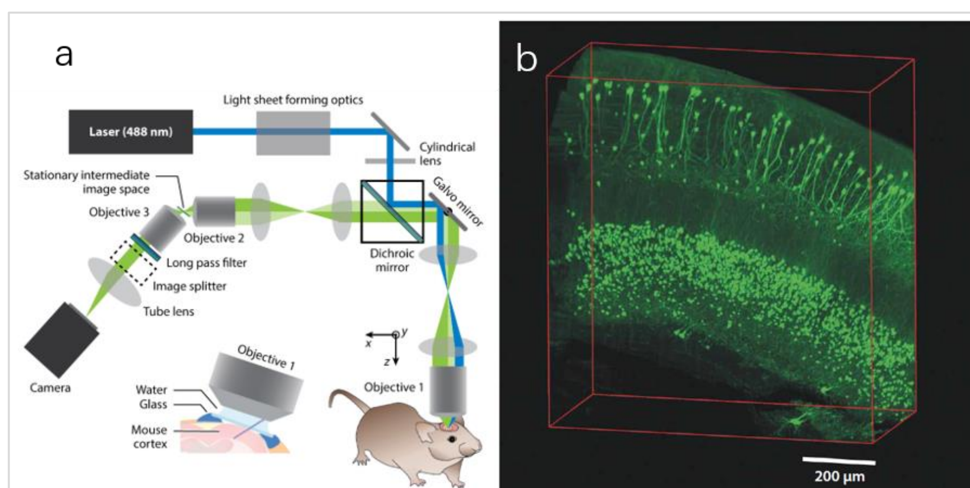


图 2.7 a)SCAPE 技术光路图 b)运用 SCAPE 技术在清醒、活动的小鼠大脑皮层第 5 层神经元的树突上的 GCaMP 活性进行成像，时间分辨率为 10VPS（图片源自[10]）。

更高的体成像速度还可以实现对以前认为棘手的样品成像。2019 年，Hillman 等人又在自由爬行的果蝇幼虫和秀丽隐杆线虫进行神经系统成像[11]。这些模式动物的活动范围都处在显微镜的视场内，可同时跟踪运动轨迹和记录神经活动。该技术不需要跟踪样品、调整 Z 轴，且没有显著光漂白。单物镜光片技术（如 SCAPE）为新一代神经科学提供了更快的成像速率，不仅允许（并记录）活体样本的 3D 运动，而且可以在实验中对运动进行测量，该技术也可用于实时跟踪记录整个样品中的细胞功能信号。

2.1.4 自由行为动物高分辨率成像

前文提到的成像技术进步可以帮助研究者在大脑多个区域的更大体积、更深深度上对神经元和突触等亚细胞结构实现更高速的成像。但是这些方法都是针对固定的活体样本或者小尺度的活动生物样本进行成像。要真正证明基于观察推断的机制，还需要对活动的、不受限制的生物(如小鼠)进行长时间活体成像。解决这个问题的一个方案是发展多光子显微技术的小型化、微型化技术。

2017 年，由北京大学分子医学研究所牵头开发的超高时空分辨微型化双光子在体显微成像系统（fast, high-resolution, miniaturized two-photon microscope, FHIRM-TPM）解决了这个问题，并在首次获取了小鼠在自由行为过程中大脑神经元和神经突触活动清晰、稳定的图像[12]。



图 2.8　FIRM-TPM 以及实验动物的实物图。（图片来自北京大学陈良怡实验室）

超高时间分辨微型化双光子荧光显微镜重量仅为 2.2 克，适于佩戴小动物头部颅窗上（图 2.8），能够实时记录数十个神经元、上千个神经突触的动态信号。在大型动物上，还可实现多探头佩戴、多颅窗不同脑区的长时程观测。该显微镜采用双轴对称高速微机电系统转镜扫描和柔性光纤，可实现 920nm 飞秒激光传输，保证了成像帧频能达到 40 赫兹（256*256 像素），同时具备多区域随机扫描和每秒 1 万线的线扫描能力。并且该显微镜实现了 0.65um 的横向分辨率，从而具有细胞与亚细胞分辨能力的活体三维成像能力。该显微镜可以观察动物在自然行为或者学习行为中多尺度、多层次神经活动的动态变化，从而观察到大脑学习、记忆、决策的过程，也可用于研究脑疾病的神经机制。

2020 年，该微型双光子显微镜系统的性能得到进一步升级（FHIRM-TPM 2.0）[13]。 通过重新设计微型光学系统，显微镜的视场（FOV）扩大了 7.8 倍，成像深度也扩展至 1mm。通过增加一个可拆卸的 Z 轴扫描模块，FHIRM-TPM 2.0 实现了体成像和多平面成像。此外，研发团队还设计了可重复装卸的头戴装置，这一装置可以在数个星期内持续记录神经元活动。该显微镜所具备的高精度、高帧率、深层成像与超长时间成像能力将改变研究者在自由活动动物中观察细胞和亚细胞结构的方式。

2.2 探针技术在神经科学领域的应用

光学成像技术的质量提升离不开新型荧光探针的发展。荧光成像需要高亮度、高特异性和高稳定的荧光标记探针。按类型荧光探针可分为合成荧光探针和遗传编码荧光探针，标记的对象包括胞内钙离子，神经递质及神经调节剂。本节将分别介绍两类探针的发展与改进。

2.2.1 化学合成的荧光探针

2019 年,Landry 等人开发了一类合成的纳米探针,这类纳米探针为近红外(1000-1300 nm)荧光探针。他们成功开发出能够在活体中标记 5-羟色胺和多巴胺的近红外纳米探针[14, 15]。5-羟色胺和多巴胺都是重要的神经调质，在调节学习记忆、注意力、情绪、睡眠和食欲中起着关键作用，其信号异常被认为是多种精神疾病的发病基础。通常的神经递质的信号传导过程位于突触间隙，然而诸如 5-羟色胺和多巴胺等神经调质可以扩散到突触外进行信号传导，使得少数神经元能够通过释放神经调质从而影响更广泛的神经元网络。这些神经调质调控的空间分布就非常值得人们关注。

Landry 等人开发的荧光探针能够指示大脑细胞外空间中调节性神经递质的浓度。他们观测到了电刺激和光遗传诱发的多巴胺在脑组织中的释放。这类探针性能稳定，且灵敏度高，特异性强。并且他们设计了了能够批量开发此类探针的方法。他们开发了一个能够富集高特异性探针的通用平台，能够针对特定的标记对象在单壁碳纳米管表面上富集出标记能力最强的探针。因此可用于开发其他神经相关的合成探针。这类纳米探针可以用作监测神经调节的多功能合成光学工具，对于观测神经调质的释放可以达到很高的空间分辨率，且开发方便，因此具有很广阔的应用前景。

2.2.2 遗传编码的荧光探针

除化学合成外，设计新的荧光蛋白分子并利用遗传手段在细胞内自行合成的遗传编码荧光探针技术也在不断发展。在钙成像探针方面，2020 年，Na Ji 团 队 通过两个单体近红外荧光蛋白，设计出了一种基于近红外荧光共振能量转移（Förster resonance energy transfer, FRET）

的基因编码钙指示剂（genetically encoded calcium indicator, iGECI）[16]。iGECI 对钙的荧光响应强度提高了 6 倍。

尽管已经被广泛运用，当前的钙荧光指示剂无法与其他探针同时使用。iGECI 则实现了生物探针和光遗传学工具的多路复用而不会相互干扰，Ji 团队验证了 iGECI 在脑切片中约 400μm 深度下成像的能力。该探针应用于混合光声荧光显微镜，可以透过完整的颅骨同时监测到小鼠大脑中神经元活动和血液的动态变化，并达到了 3μm 左右的横向分辨率和 25-50μm 左右的轴向分辨率。而在双光子成像中，他们在小鼠视觉皮层中检测到了诱发及自发的神经元活动。iGECI 有望成为一种非常实用的新型钙成像工具。

在神经递质方面，人们开发了基于 G 蛋白偶联受体（GPCR）的探针，从而以高时空分辨率观测活动生物的神经化学物质的动态变化。2020 年，李毓龙团队改进了他们在 2018 年开发的用于观测乙酰胆碱的探针 $GRAB_{ACh}$（GPCR-activation-based Ach）[17, 18]，乙酰胆碱是一种重要的神经递质。 2018 年开发的第一代探针已经具备亚秒级的时间响应和高生理亲和力。新一代探针更兼具了特异性、高灵敏性、高亲和力、快速反应速率以及高选择性。该探针还能够对不同行为下的小鼠的多脑区的乙酰胆碱活动进行动态检测。

李毓龙团队在 2020 年还开发了新型的具有红色荧光的多巴胺探针 $GRAB_{DA}$，并优化了绿色荧光 $GRAB_{DA}$ 探针，在反应幅度上提升了 2-3 倍[19, 20]。这样就可以在动物体内同时表达红色的 $GRAB_{DA}$ 和其他绿色荧光探针（如钙离子探针，神经递质探针等），实现多种信号的同时记录。多巴胺和乙酰胆碱等神经递质在大脑中有至关重要的作用，直接以高时空分辨率观测这些神经递质的动态变化对于了解动物生理功能与神经活动有着重要的意义。因此基于 GPCR 的神经化学成像探针在未来将有广泛应用。

2.3 神经调控技术在认知科学领域的应用

2.3.1 个体化非侵入式神经调控技术应用于大脑视觉认知功能的增强

基于现有技术，如果要对视觉功能正常的健康人群进一步增强其视觉认知功能，一般采用非侵入式的电磁调控手段以规避植入物带来的风险。例如，经颅电刺激（transcranial electrical stimulation, tES）用于增强视觉运动学习和减弱视觉拥挤效应等。然而，由于经颅电磁刺激需要透过头皮、颅骨及脑脊液等层层组织和屏障，最终在大脑皮层形成电场来实现对神经系统的间接调控，刺激精度相对直接与大脑皮层接触的侵入式电刺激手段要低得多。因此，如何确定最优的刺激方案以实现精准而一致性的视觉认知功能调控效果，是研究者共同面临的挑战。接下来，我们将以两项近期的高水平研究为例，介绍未来视知觉功能增强可采用的个体化神经调控技术。

2020 年 9 月，Nicholas L. Balderston 等发表在《Nature Protocols》的研究[21]，介绍了基于功能磁共振成像(functional magnetic resonance image, fMRI)和电场优化仿真实现针对特定被试的靶向经颅磁刺激（transcranial magnetic stimulation, TMS）的详细方法和流程。简单来说，该方案需要设计 fMRI 实验在单个被试水平进行最佳 TMS 靶点定位，再基于被试结构像数据建立用于电场仿真的有限元头部模型，最后依据建模结果选择在靶点位置形成最大电流的最优线圈取向。类似地，现有的优化 tES 研究多基于高密度头皮表面电极阵列（如按照 10-10 国际标准脑电电极位置布局），以最大化刺激目标区域的电场强度或聚焦度为目标，结合算法计算每个电极位置应施加的电流强度。在刺激目标靶点定位方面，这篇文章的原理也可以推广到基于电场优化和 fMRI 导向的个体化 tES 方案的制定上，并直接应用到视知

觉功能增强的相关研究中。

除了需要结合结构像进行个体化建模、结合功能像定位精准刺激目标，个体化神经调控技术还需要考虑刺激前大脑活动状态的个体差异以确定刺激参数，而这往往通过采集被试的脑电活动来评估实现。例如，2021 年初， Robert M. G. Reinhart 课题组在 《Nature Medicine》杂志发表的文章报告[22]，基于奖励网络的个性化经颅交流（transcranial alternating current stimulation, tACS）作用于眶额皮质(orbitofrontal cortex, OFC)后，减少了强迫症的行为。研究者首先记录了被试执行强化学习任务时的脑电图活动（图 2.9a），确定被试与内在奖励网络相关的 β-γ 神经活动频率值，然后针对双侧 OFC 的内侧靶点计算刺激电极的位置及对应的刺激强度（图 2.9b），确定这些参数后在被试做任务期间使用高分辨率tACS（High-definition tACS, HD-tACS）以调节 OFC 的神经活动节律。实验表明对 OFC 进行 5 天 β-γ 频段的调控，可有效减弱非临床人群的强迫症行为达 3 个月。这项研究采用的依据特定被试在执行与金钱相关的强化学习任务（Monetary reinforcement learning task）中的脑电活动来确定 tACS 刺激频率值的方法，可类似地应用到视知觉任务中，以实现对具体被试的个体化视知觉功能调控。

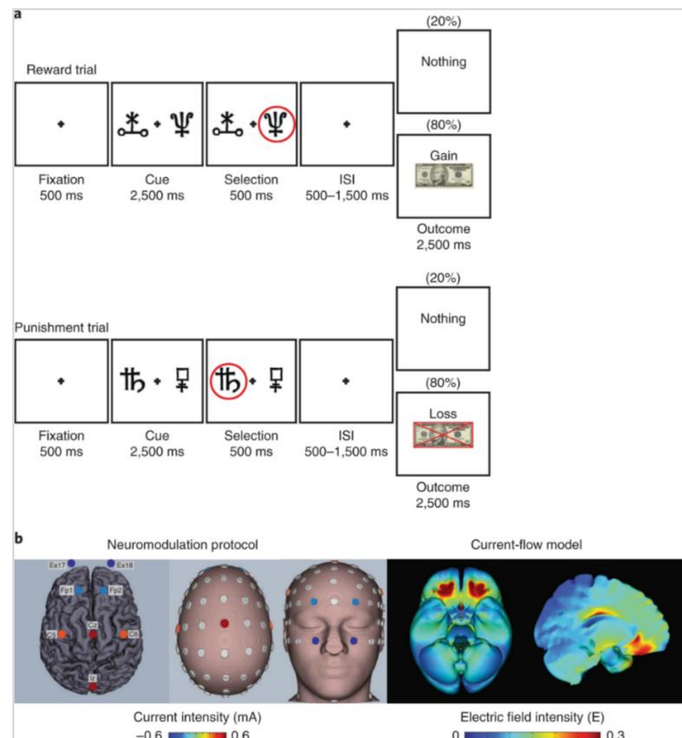


图 2.9 金钱强化学习任务和眶额神经调节方案（图片源自[22]）

2.3.2 新型侵入式神经调控技术应用于大脑视觉认知功能的重建

在早期的研究中，研究者发现通过给人类视觉皮层施加电刺激可引起短暂的光闪烁知觉（光幻视）。此后，科学家们致力于开发通过电刺激视觉皮层帮助后天失明盲人恢复视觉的视皮层假体（visual cortical prostheses, VCP）。VCP 的基本结构包括：用于捕获视觉场景的头戴式摄像头模块、将图像转换成对应刺激模式的计算模块，以及电刺激信号产生模块。近年来，随着小型化技术及高效率数字电路技术的进步，VCP 得以进一步发展。其中，2020 年荷兰神经科学研究所 Pieter R. Roelfsema 课题组和美国贝勒医学院 Daniel Yoshor课题组发表

的重要成果，从不同的技术角度利用 VCP 分别探索了猕猴[23]和人类[24]的视知觉功能增强新方案，在众多研究成果中令人耳目一新。

Pieter R. Roelfsema 等在《Science》杂志发表的研究[23]沿用了早期研究中同时刺激多个电极点时被试能知觉到对应形状的假设，借助了在生物兼容性和电学特性上表现优异的犹他电极阵列，通过在猕猴的视觉皮层中植入 16 个 64 通道的微电极阵列（图 2.10 A,B），实现了对于视觉皮层上千个位点的活动调控。以往的研究表明，当初级视皮层施加电刺激成功产生幻视知觉时，可在更高级的视皮层检测到相应的响应。基于此，两个 64 通道的犹他电极阵列被植到了 V4 区以采集高级视皮层的电生理活动，通过与剩余 14 个位于初级视皮层 V1 区的电极阵列配合，即可确定每个电刺激通道的刺激阈值。这一点与之前的常规操作不同，它不再需要由使用者针对不同通道的不同刺激水平逐个报告感觉，而是可以自动快速地进行阈值测试，这大大提升了视觉重建系统的效率和实用性。在明确每个刺激通道对应的光幻视阈值及出现的位置后（图 2.10 C-F），研究者们利用组合电刺激对猕猴刺激出了点、线及字母的光幻视知觉。在具体的实验任务中，Roelfsema 组对两只植有高密度电极阵列的猕猴进行了运动方向任务的训练和测试，与运动轨迹感受野对应的微电极被施加顺序电刺激，考察猕猴是否可以感知相继生成的光幻视点的方向。而在另外的字母识别任务中，字母形状被分解成点状并显示在计算机显示器上（图 2.11），电刺激信号被同时传递到 8～15 个电极通道上，尝试对猕猴构建出对字母形状的知觉。而最终试验结果表明，猴子能成功“看到”移动的点、线条，以及字母。

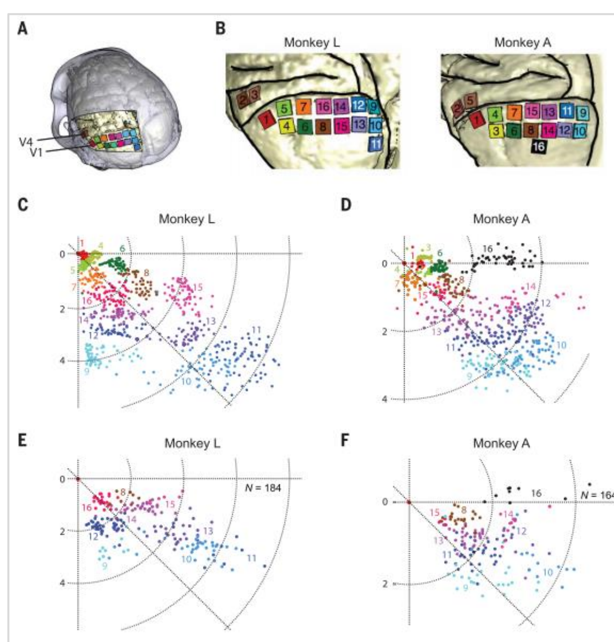


图 2.10 电极植入位置及对应的感受野（图片源自[23]）

A.实验猴 L 左半脑V1 和Ⅴ4 植入电极位置示意图。B.实验猴 L 和 A 植入电极位置的放大示意图。C.实验猴 L 颅内电极对应的感受野中心。D.实验猴 A 颅内电极对应的感受野中心。E.实验猴 L 颅内电极对应的光幻视位置。F.实验后 A 颅内电极对应的光幻视位置。

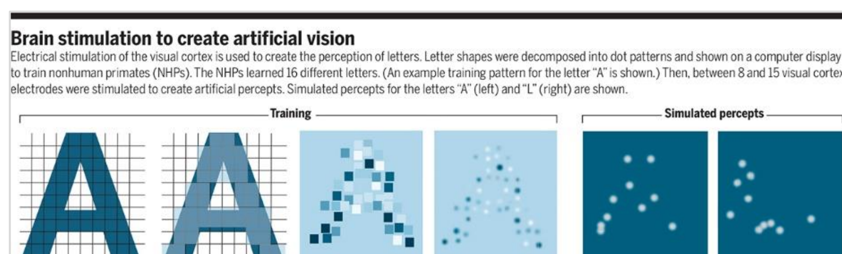


图 2.11 视觉皮层的电刺激用于重建对字母的知觉（图片源自[23]）

如果说高密度电极阵列的成功植入是保障 Roelfsema 等研究顺利的关键因素，那么美国贝勒医学院 Daniel Yoshor 课题组在《Cell》杂志发表的关于利用动态颅内电刺激新技术（dynamic current steering）的研究[24]，则是从刺激范式上实现了创新，该研究利用了较少的电极成功地在人类（三名癫痫患者以及两名盲人）初级视皮层“打印”出字母形状。据报道，这项研究的灵感来源于一个简单的触觉实验：当使用多个探针按压手掌形成字母形状时，大脑无法感知刺激的意义、无法辨别目标字母的形状，而当使用单个探针以与目标形状相匹配的顺序在手掌上动态地示踪时，大脑却可以连贯地知觉到字母形状（图 2.12A,B）。类似地，在过去的研究中，通过简单地对多个刺激通道同时施加电流（图 2.12C），只会让被试知觉到多个单独的光幻视点而难以整合成有意义的形状，并且这些独立的光幻视点可能会混在一起形成一个单独的幻视点难以被分辨。与此不同，Yoshor 等开发的动态电刺激技术则是通过控制各通道电流以与目标形状（图 2.12D 白色箭头）匹配的时间顺序先后传递到电极网格，从而产生连贯的视觉感知。在这个过程中，可以通过调整相邻两个电极注入电流的强度（图 2.13），以激活这两个电极中间位置的神经元，从而在这两个单独电极（图 2.12D 深蓝色圆圈）之间的位置达到虚拟电极（图 2.12D 浅蓝色椭圆形）电刺激的效果，实现刺激位置的连贯性。在 Yoshor 课题组的这项研究中，研究者对初级视皮层植有微电极阵列的癫痫病人和盲人测试了这项技术，结果显示（图 2.14），通过在对应的视觉皮层区域施加字母形状的动态电刺激，患者可以“看到”电刺激所示踪的形状，并在触摸屏上准确地“打印”出这些形状（如“W”，“S”和“Z”等字母）。不仅如此，这项技术还表现出较高的刺激效率，被试可以快速识别和重建形状，其效率可达每分钟“打印”86 个形状。值得注意的是，Yoshor组研究涉及的参与者无法通过静态电刺激识别任何形状，这与早期研究声称同时对多通道施加静态电刺激能够使参与者产生形状的知觉不太一致。

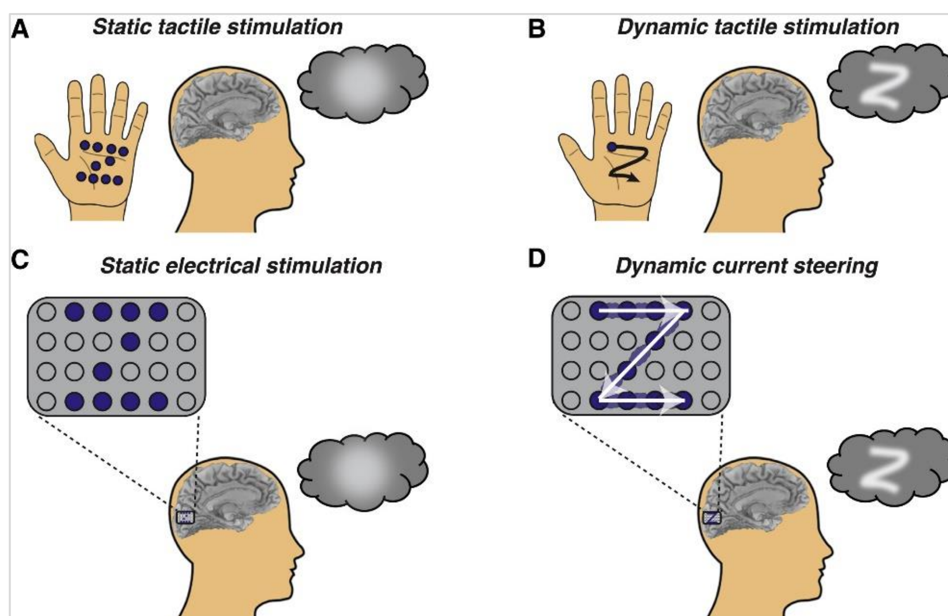


图 2.1 动态刺激重建触觉和视觉功能的示意图（图片源自[24]）

A．大脑无法感知静态触觉刺激对应的字母形状。B.动态触觉刺激示踪字母形状后大脑感知到对应形状。C.多通道颅内电极稳态电刺激无法使大脑知觉出字母的形状。D.动态颅内电刺激实现大脑对字母形状的视知觉。

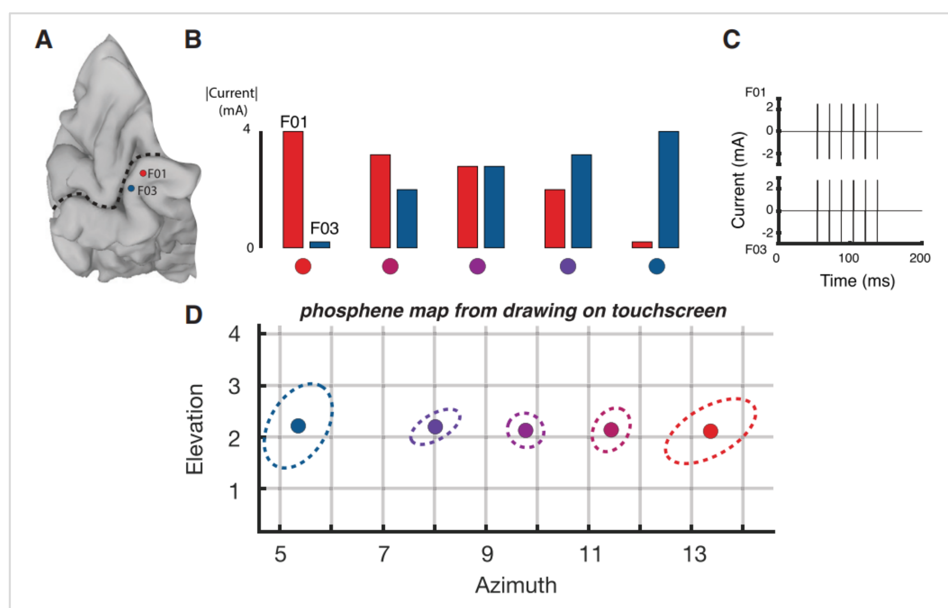


图 2.13 通过调整相邻电极刺激强度实现中间位置的虚拟电刺激（图片源自[24]）

A.相邻刺激位点（F01 和 F03）植入位置示意图。B.F01 和 F03 刺激强度调节。

C.刺激电流波形示意图。D.图 B 中调节 F01 和 F03 刺激强度对应形成的光幻视位置。

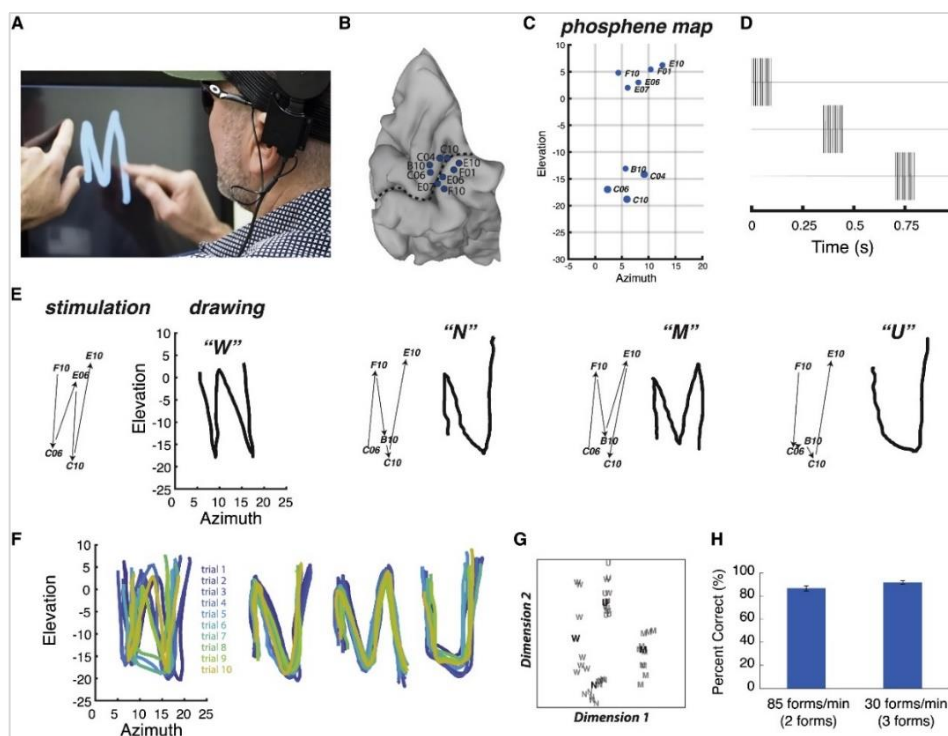


图 2.14 对盲人测试动态刺激（图片源自[24]）

A. 实验录像截图。B.电极植入位置示意图。C.光幻视位置示意图。D.电刺激波形示意图。

E.电刺激次序与患者对应绘出的形状。F.患者在多个试次中绘制的形状。

G.对患者在每个试次绘出的形状的多维尺度分析。H.字母形状识别准确率。

可以看到,对于同时刺激多个电极位点是否会引起可分辨的视知觉,Roelfsema 等与Yoshor 等在利用视皮层刺激实现诸如字母形状等有意义的幻视时，持有不同的观点，且在二者分别针对猕猴和人类的实验中，得到的结果也具有不一致性。此外，Roelfsema 等认为Yoshor 等提出的动态电刺激方法难以保证单位时间内传达的信息总量,这一点也许在未来超高密度微电极阵列具备植入人类大脑的条件时能得到解决。

延伸阅读：动态颅内电刺激应用于高级视觉皮层产生更真实的视知觉

上述最前沿的视皮层假体研究已经出色地解决了如何呈现简单无色平面形状的问题。然而，视觉世界充斥着具有彩色、立体等信息的自然场景。未来的研究趋势应该会从如何电刺激产生带有彩色、立体信息的光幻视入手，从而使 VCP 足以产生更真实的视知觉。在这里，我们要介绍一下 2019 年分别发表在《Science》和《Cell》的两项研究[25, 26]，探讨高密度电极阵列及动态颅内电刺激应用于高级视觉皮层产生信息更丰富的视知觉的新方向。这两篇文章结合颅内神经信号记录，采用计算神经科学的智能算法和神经网络生成“新”图像，实现了对高级视皮层神经元的选择性激活。

在 James J. DiCarlo 课题组于 2019 年 5 月在《Science》杂志发表的研究中[25]，三只猕猴的 V4 区分别被植入了 96 通道的犹他电极阵列，通过结合神经电生理活动与人工神经网络模型，研究计算机合成的“新”图像调控神经元活动的可能性[25]。首先使用自然图像集训练特定的深度人工神经网络以得到稳定和优化的网络参数，然后结合颅内神经信号将这个神经网络的 L3 层神经元映射到猕猴 V4 区的记录位点，并利用该层神经元模型的响应优化刺激图像，以针对每个记录位点合成对应的发放偏好图像（称为合成“控制器”图像），最

后将这些合成的图像刺激与自然图像刺激分别呈现给猕猴，比较 V4 区的各神经信号记录位点的发放情况（图 2.15）。实验考察了自模型“神经元”与神经信号采集位点映射到测试期间信号有效的电极通道，三只猕猴总共有 109 个可靠的记录位点，对应于深度人工神经网路 L3 层的 109 个模型“神经元”。实验结果显示，这些记录位点对合成的“控制器”图像相对于自然图像而言表现出更高的发放率（提高 57%），且在选择合适的图像合成原则后神经元发放呈现出一定的选择性特征，即合成图像显著激活特定的记录位点（图 2.16）。此外，对于感受野高度重叠的多个记录位点，可合成系列“控制器”图像对它们实现单独控制，如图 2.17 所示八个不同的刺激图片依次最大化某位点的发放率而抑制其他位点的神经响应。从结果上看，这些可以有选择地激活 V4 区神经元的图像，带有除形状信息外更加高级的纹理信息。

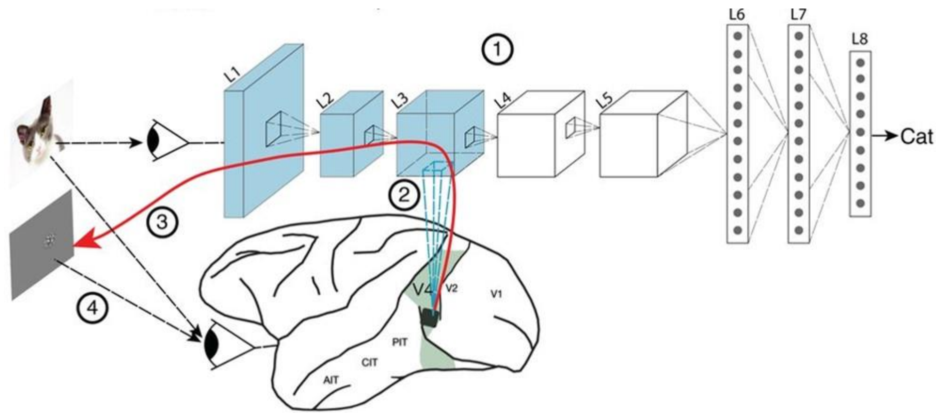


图 2.15 合成图像实现神经控制的实验流程图（图片源自[25]）

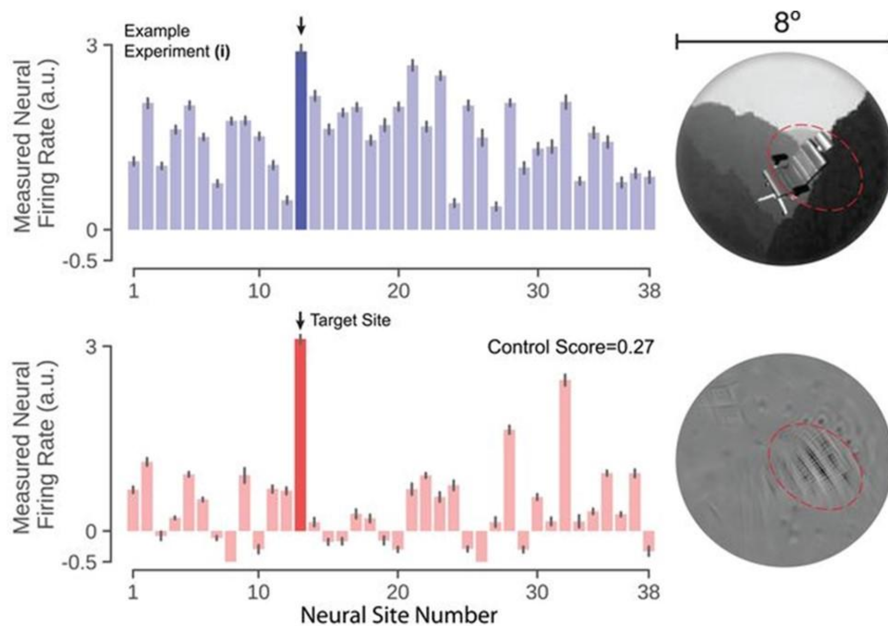


图 2.16 自然图像（第一行）与合成图像（第二行）刺激下记录位点的激活情况

（图片源自[25]）

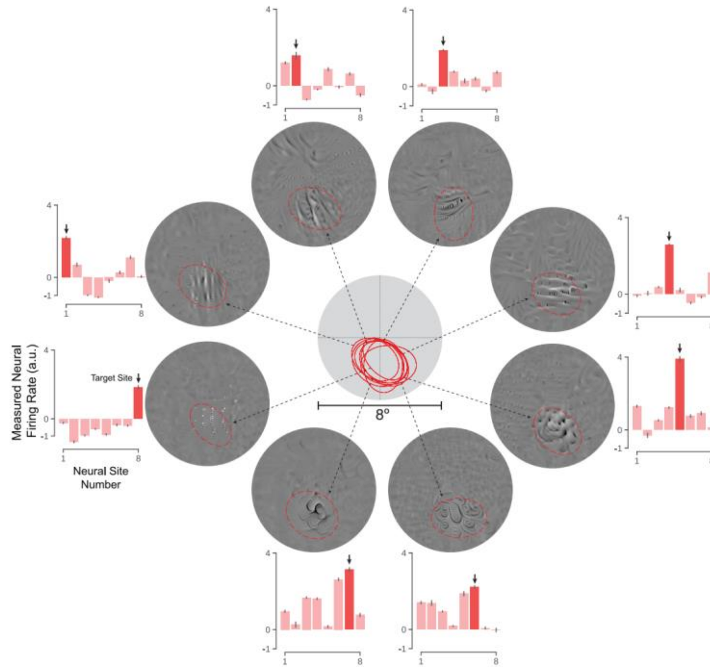


图 2.17 针对感受野高度重叠的八个记录位点的选择性控制（图片源自[25]）

相对而言，Margaret S. Livingstone 课题组发表在《Cell》上的研究[26]则是采用的一种类似于在线神经反馈的方式，基于神经元的发放特点逐步演变生成可以显著提高神经元兴奋性的新图像。其中，神经电极阵列位于猕猴的颞下（inferotemporal，IT）皮质。以随机的黑白纹理图像为起点，基于 IT 区神经元的实时响应评估偏好程度，然后结合深度生成神经网络与遗传算法生成新的图像（图 2.18），经过多次演变后最终生成能大幅提高神经元活性的图像（图 2.19A）。图 2.19B显示了由 XDREAM（EXtending DeepDream with Real-time Evolution for Activity Maximization in real neurons）生成的可诱发某神经元强烈发放的图像，该实验猕猴的这个神经元对自然图像集中的躯干和猕猴面孔反应较强（图 2.19C和图 2.19D 分别为该神经元响应最大和最小的十张图像）。以往的研究大多基于已知的、与真实世界相关的图像测试神经元的偏好，但这带有很强的前提特征和假设。而 Livingstone 等的这项研究类似于将“作画”的主动权交给了神经元，由神经元响应决定图像的演变进程，更能接近研究问题的本质。而结果显示，IT 神经元“画”出的图形带有丰富的颜色、形状和纹理信息。

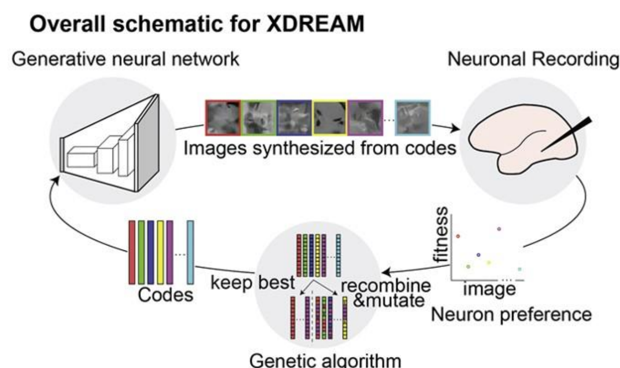


图 2.18 XDREAM 算法的整体框架图（图片源自[26]）

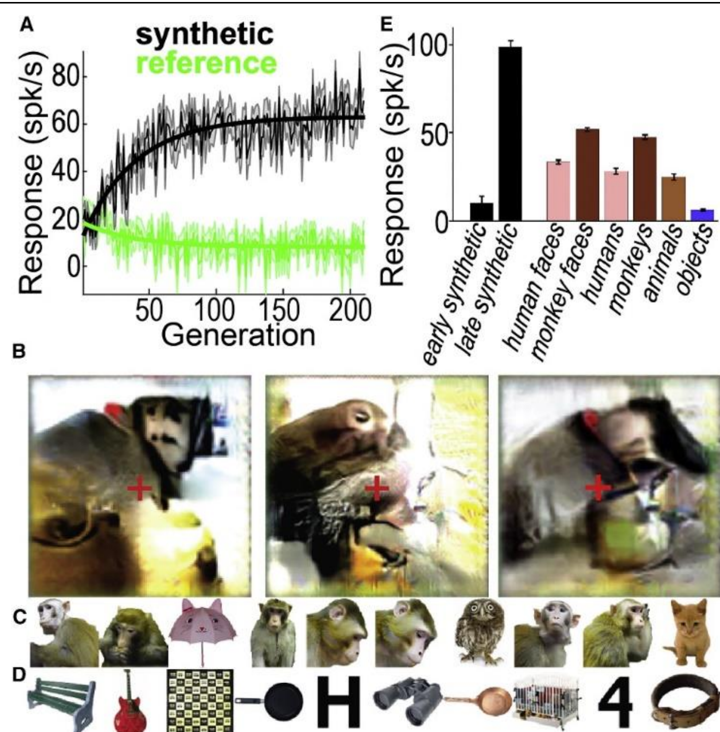


图 2.19 生成的图像和自然图像以及它们诱发的神经元反应（图片源自[26]）

A. 神经元分别对合成图片（黑色）和参考图片（绿色）的平均响应。B. 采用不同的方法合成的最终图像。左图对应 A 中的黑色发放。C. 神经元发放最强的 10 + 张自然图像。D. 神经元发放最弱的十张图像。E. 神经元对不同类型图像的选择性响应。

上面介绍的两项研究均采用了颅内神经元活动采集技术，分别与机器学习的判别模型和生成模型结合，生成了可相对选择性地提高神经记录位点或单个神经元发放的图像刺激。并且，随着神经网络及深度学习模型的进一步完善，利用这种思路，将实现更精准的神经元级别的操控，而且这有助于理解高级视皮层对带有纹理、颜色等复杂视觉信息的处理机制。在这里我们大胆设问，如果将高密度电极阵列植入人体高级视觉皮层，依据神经网络生成的图片刺激及高级视皮层神经元发放的对应关系，采用新型动态颅内电刺激技术作用于高级视皮层，能否形成更接近于自然场景的视知觉？这种新型视觉认知功能调控手段值得认知神经科学和计算神经科学领域的研究者共同进行深入的探讨，这将在神经科学的研究领域和相关医疗应用领域产生重大影响。

2.4 总结与展望

在本部分，我们总结了近一两年来认知与神经科学在技术方面的一些进展与突破，重点介绍了神经科学领域相关成像技术上的进步，以及认知功能领域神经调控技术应用于视觉功能的增强与重建层面的进展。

成像技术方面，我们介绍了多光子显微技术、宽场显微技术和光片显微技术等显微成像技术在深层成像，超高速成像、高精度成像、超大视场成像和微型化等方向的进展。同时还介绍了成像相关工具的发展，包括图像数据的处理工具与处理流程的改进，以及荧光标记分子工具的发展。面向未来，我们认为系统神经科学的研究将朝着对生物大脑实现更大探测范围和更高时空分辨率的方向发展。具体到成像领域，可以预见会如下的发展趋势：

（1）更深层的探测能力。当前成像技术已经能达到毫米级的成像深度，未来技术将向着能够在复杂活体组织中实现厘米级深层成像的方向努力。新型光源的开发，多种探测手段结合（如光声成像）或者红外与双光子成像技术的改进将有助于深层成像的发展。

（2）更高的时空分辨率。当前成像技术在空间分辨率上已经达到亚细胞水平，采样频率也能达千赫兹。但是开发同时具备大尺度，高时空分辨率的体成像技术依然面临很多困难。高时空分辨率的三维体成像带来的大规模数据集（至少 1TB/h 的记录速度）的处理也会是非常大的挑战。实现这一目标在硬件上有赖于探针技术和成像方法的进步。在数据处理上则需要开发一系列自动化的分析工具，尤其是结合时间序列的 4D 去噪算法和 3D 跟踪、配准、信号提取以及结果可视化等数据处理工具的开发。深度学习方法在成像数据中的运用有希望实现这些分析工具。

在神经调控技术方面，更强大的计算模型和工程技术实现将为科学理论转化为具有实用价值的调控手段提供强有力的推动。在本部分非侵入式电磁刺激调控大脑视觉认知功能的例子中，结合被试的大脑状态、认知水平、大脑解剖结构及功能区域定位等个体化信息定制刺激方案的方法流程，可有效的提高调控的准确性、有效性和一致性，未来相关的研究将致力于提高这种流程的自动化程度，并将这种方法应用到多种实验中以测试效果，进行迭代改进。而在侵入式方法实现视觉重建方面，一方面人们期待未来以柔性电极为代表的高通量、高生物相容性电极技术能够赋予人们在更长时间（数年甚至更长）中以更高的精度（数万通道甚至更多）进行视觉重建刺激，一方面对于视觉编码的进一步深入理解，以及充分结合人工智能技术的最新发展，将有望产生更加高效的刺激编码算法。这些方向的进展，将大大促进新型电刺激手段产生更丰富、更真实的视觉感知，最终对受损的视觉功能实现替代。

参考文献

[1] Augustinaite, S. and B. Kuhn. 2020. "Complementary Ca(2+) Activity of Sensory Activated and Suppressed Layer 6 Corticothalamic Neurons Reflects Behavioral State." Curr Biol 30 (20): 3945-3960 e5.

[2] "Movable Objective Microscope®." https://www.sutter.com/MICROSCOPES/mom.html.

[3] "Imaging brains of mice on the go." https://www.nature.com/articles/d42473-020-00388-2#ref-CR1.

[4] Chow, D.M., D. Sinefeld, K.E. Kolkman, D.G. Ouzounov, N. Akbari, R. Tatarsky, A. Bass, C. Xu and J.R. Fetcho. 2020. "Deep three-photon imaging of the brain in intact adult zebrafish." Nat Methods 17 (6): 605-608.

[5] Hinsdale, T., B.H. Malik, C. Olsovsky, J.A. Jo and K.C. Maitland. 2015. "Volumetric structured illumination microscopy enabled by a tunable-focus lens." Opt Lett 40 (21): 4943-6.

[6] Shi, R., C. Jin, H. Xie, Y. Zhang, X. Li, Q. Dai and L. Kong. 2019. "Multi-plane, wide-field fluorescent microscopy for biodynamic imaging in vivo." Biomed Opt Express 10 (12): 6625-6635.

[7] Wu, J., Y. Liang, S. Chen, C.L. Hsu, M. Chavarha, S.W. Evans, D. Shi, M.Z. Lin, K.K. Tsia and N. Ji. 2020. "Kilohertz two-photon fluorescence microscopy imaging of neural activity in vivo." Nat Methods 17 (3): 287-290.

[8] Fan, J.L., J.A. Rivera, W. Sun, J. Peterson, H. Haeberle, S. Rubin and N. Ji. 2020. "High-speed volumetric two-photon fluorescence imaging of neurovascular dynamics." Nat Commun 11 (1): 6020.

[9] Bouchard, M.B., V. Voleti, C.S. Mendes, C. Lacefield, W.B. Grueber, R.S. Mann, R.M. Bruno and E.M. Hillman. 2015. "Swept confocally-aligned planar excitation (SCAPE) microscopy for high speed volumetric imaging of behaving organisms." Nat Photonics 9 (2): 113-119.

[10] Hillman, E.M.C., V. Voleti, K. Patel, W.Z. Li, H. Yu, C. Perez-Campos, S.E. Benezra, R.M. Bruno and P.T. Galwaduge. 2018. "High-speed 3D imaging of cellular activity in the brain using axially-extended beams and light sheets." Current Opinion in Neurobiology 50: 190-200.

[11] Vaadia, R.D., W.Z. Li, V. Voleti, A. Singhania, E.M.C. Hillman and W.B. Grueber. 2019. "Characterization of Proprioceptive System Dynamics in Behaving Drosophila Larvae Using High-Speed Volumetric Microscopy." Current Biology 29 (6): 935-+.

[12] Zong, W., R. Wu, M. Li, Y. Hu, Y. Li, J. Li, H. Rong, H. Wu, Y. Xu, Y. Lu, H. Jia, M. Fan, Z. Zhou, Y. Zhang, A. Wang, L. Chen and H. Cheng. 2017. "Fast high-resolution miniature two-photon microscopy for brain imaging in freely behaving mice." Nat Methods 14 (7): 713-719.

[13] Zong, W., R. Wu, S. Chen, J. Wu, H. Wang, Z. Zhao, G. Chen, R. Tu, D. Wu, Y. Hu, Y. Xu, Y. Wang, Z. Duan, H. Wu, Y. Zhang, J. Zhang, A. Wang, L. Chen and H. Cheng. 2021. "Miniature two-photon microscopy for enlarged field-of-view, multi-plane and long-term brain imaging." Nat Methods 18 (1): 46-49.

[14] Jeong, S., D. Yang, A.G. Beyene, J.T. Del Bonis-O'Donnell, A.M.M. Gest, N. Navarro, X. Sun and M.P. Landry. 2019. "High-throughput evolution of near-infrared serotonin nanosensors." Sci Adv 5 (12): eaay3771.

[15] Beyene, A.G., K. Delevich, J.T. Del Bonis-O'Donnell, D.J. Piekarski, W.C. Lin, A.W. Thomas, S.J. Yang, P. Kosillo, D. Yang, G.S. Prounis, L. Wilbrecht and M.P. Landry. 2019. "Imaging striatal dopamine release using a nongenetically encoded near infrared fluorescent catecholamine nanosensor." Science Advances 5 (7).

[16] Shemetov, A.A., M.V. Monakhov, Q.R. Zhang, J.E. Canton-Josh, M. Kumar, M.M. Chen, M.E. Matlashov, X. Li, W. Yang, L.M. Nie, D.M. Shcherbakova, Y. Kozorovitskiy, J.J. Yao, N. Ji and V.V. Verkhusha. 2020. "A near-infrared genetically encoded calcium indicator for in vivo imaging." Nature Biotechnology.
[17] Jing, M., Y. Li, J. Zeng, P. Huang, M. Skirzewski, O. Kljakic, W. Peng, T. Qian, K. Tan, J. Zou, S. Trinh, R. Wu, S. Zhang, S. Pan, S.A. Hires, M. Xu, H. Li, L.M. Saksida, V.F. Prado, T.J. Bussey, M.A.M. Prado, L. Chen, H. Cheng and Y. Li. 2020. "An optimized acetylcholine sensor for monitoring in vivo cholinergic activity." Nat Methods 17 (11): 1139-1146.
[18] Jing, M., P. Zhang, G. Wang, J. Feng, L. Mesik, J. Zeng, H. Jiang, S. Wang, J.C. Looby, N.A. Guagliardo, L.W. Langma, J. Lu, Y. Zuo, D.A. Talmage, L.W. Role, P.Q. Barrett, L.I. Zhang, M. Luo, Y. Song, J.J. Zhu and Y. Li. 2018. "A genetically encoded fluorescent acetylcholine indicator for in vitro and in vivo studies." Nat Biotechnol 36 (8): 726-737.
[19] Sun, F., J. Zhou, B. Dai, T. Qian, J. Zeng, X. Li, Y. Zhuo, Y. Zhang, K. Tan and J. Feng. 2020. "New and improved GRAB fluorescent sensors for monitoring dopaminergic activity in vivo." BioRxiv.
[20] Sun, F., J. Zhou, B. Dai, T. Qian, J. Zeng, X. Li, Y. Zhuo, Y. Zhang, Y. Wang, C. Qian, K. Tan, J. Feng, H. Dong, D. Lin, G. Cui and Y. Li. 2020. "Next-generation GRAB sensors for monitoring dopaminergic activity in vivo." Nat Methods 17 (11): 1156-1166.
[21] Balderston, N.L., C. Roberts, E.M. Beydler, Z.D. Deng, T. Radman, B. Luber, S.H. Lisanby, M. Ernst and C. Grillon. 2020. "A generalized workflow for conducting electric field-optimized, fMRI-guided, transcranial magnetic stimulation." Nat Protoc 15 (11): 3595-3614.
[22] Grover, S., J.A. Nguyen, V. Viswanathan and R.M.G. Reinhart. 2021. "High-frequency neuromodulation improves obsessive-compulsive behavior." Nat Med 27 (2): 232-238.
[23] Chen, X., F. Wang, E. Fernandez and P.R. Roelfsema. 2020. "Shape perception via a high-channel-count neuroprosthesis in monkey visual cortex." Science 370 (6521): 1191-1196.
[24] Beauchamp, M.S., D. Oswalt, P. Sun, B.L. Foster, J.F. Magnotti, S. Niketeghad, N. Pouratian, W.H. Bosking and D. Yoshor. 2020. "Dynamic Stimulation of Visual Cortex Produces Form Vision in Sighted and Blind Humans." Cell 181 (4): 774-783 e5.
[25] Bashivan, P., K. Kar and J.J. DiCarlo. 2019. "Neural population control via deep image synthesis." Science 364 (6439).
[26] Ponce, C.R., W. Xiao, P.F. Schade, T.S. Hartmann, G. Kreiman and M.S. Livingstone. 2019. "Evolving Images for Visual Neurons Using a Deep Generative Network Reveals Coding Principles and Neuronal Preferences." Cell 177 (4): 999-+.

第三部分 神经科学对人工智能的启发

大脑是目前已知的宇宙中唯一实现了通用智能的信息加工载体。理解大脑的工作机制对我们发展通用人工智能具有重要的启示作用。大脑是亿万年大自然进化的产物，其通过不断地演化、试错、选择，获取了有效的智能信息加工的优化结构，包括功能分区、网络的优化拓扑结构、信息的优化表征、以及信号的高效传输等。这些先天结构通过遗传方式保留下来，使得生物体可以很好地生存于外部世界，并且学会适应不断变化中的外部世界。这些先天结构是大脑信息处理的高效性、鲁棒性、和节能性等优点的神经基础。因此，从大脑中寻找和归纳出这些先验知识，为人工智能的快速发展提供了一种有效的途径。神经科学，尤其是计算神经科学，在这个借鉴过程中扮演了重要的角色，其对实验发现进行归纳和总结，抽取出大量实验数据和现象背后的计算规则和原理，并通过数学建模进行验证。目前，这些归纳出的计算原理和模型正逐步被用于指导人工智能算法的设计，我们预计在未来几年，这类受神经科学启发的人工智能研究会越来越多。本章总结了 2020 年四项受神经科学启发的人工智能算法和应用，包括借鉴秀丽线虫的神经网络模型实现优化的运动控制，借鉴果蝇的神经网络模型进行自然语言处理，以及模拟生物初级视觉系统的结构特点或者群编码特征来提高深度学习网络对噪声的鲁棒性。

3.1 秀丽线虫的神经网络用于运动控制

秀丽线虫是一种非常初等的动物，但经过漫长的时间进化，其也形成了较为优化的神经网络系统和信息处理机制。秀丽线虫尽管只有 300 多个神经元，却可以实现蠕动，运动控制以及导航。Mathias Lechner 等人借鉴了秀丽线虫的神经系统[1],构建了一个神经环路策略(NCPs, neural circuit polices)，其与卷积神经网络结合，构成了一个端到端的学习系统。研究发现，此系统只有 19 个控制神经元，通过 253 个突触将 32 个封装的输入特征连接到输出，尽管只有少量的参数，也可以学习把高维输入映射到操纵命令。这一系统被成功地应用于特定场景的自动驾驶任务，相比于传统的方法，表现出更好的鲁棒性以及可解释性。

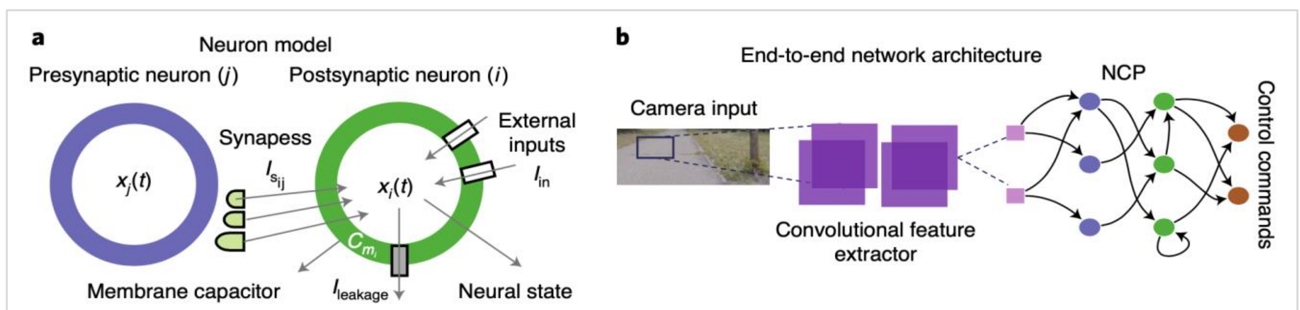


图 3.1 基于 LTC 神经元模型设计 NCP 网络，图片源自[1]。

在秀丽线虫的神经系统中，很多神经回路均由不同的四层拓扑网络结构组成，这一网络结构接受外部的感知输入，然后传送至中间神经元，命令神经元，最后产生输出并传送至运动神经元，用于控制产生最后的肌肉运动。在秀丽线虫的神经系统中，这一拓扑连接结构具有 90% 的稀疏度。并在中间神经元之间和命令神经元之间存在大量互馈连接。这一连接结构具备很多诱人的计算性质，如高效地分布式控制，层级化的时间动力学等。Mathias Lechner 等人借鉴了秀丽线虫这一拓扑结构，构建了 NCPs，用于建模时序信息。人工神经元通常采用一个简单的非线性函数来表示，而 NCPs 中的每一个仿生神经元均是一个一维的微分动力系统，每个动力系统的时间常数可随输入而改变。在自动驾驶任务中，算法通过一个摄像头接受连续的外部图像输入，经卷积神经网络提取特征，送入 NCPs 中进行整合，输出控制信号，控制方向盘转动，

如图 3.1 所示。训练数据是人类的驾驶视频，包括视频图像和相应的方向盘转动关联数据。实验发现相比于卷积神经网络加上传统互馈网络（如 LSTM，GRU 等）组成的端到端学习系统，卷积神经网络加上 NCPs 具备更好的鲁棒性和可解释性。当图像输入加入噪声时，NCPs 抗干扰性比 LSTM 等更强。NCPs 只有 19 个神经元组成，通过可视化，发现训练完成后的网络，有的神经元专注于道路边界检测，有的专注于视野中的地平线检测等。NCPs 具有更少的参数量，所需要的训练时间相比于 LSTM 等也更短。

3.2 果蝇的神经网络用于自然语言处理

果蝇的大脑已被神经科学家广泛地研究。果蝇的大脑中有一部分叫蘑菇体，其可以用于处理气味，视觉，温度以及湿度等信息输入，比如区分偏好刺激和危险刺激，从而帮助果蝇趋利避害。蘑菇体结构已经被详细解析了，其包含 2000 个 Kenyon 细胞，这些系统会接受多模态的感知觉输入，如图 3.2 所示。这些 Kenyon 神经元同时会接受一个前对外侧神经元的抑制性输（anterior paired lateral neuron），从而外界输入可以在 Kenyon 细胞群中形成一个高维的、稀疏的表征。那么类似蘑菇体的网络结构是否可以用于解决其它任务？

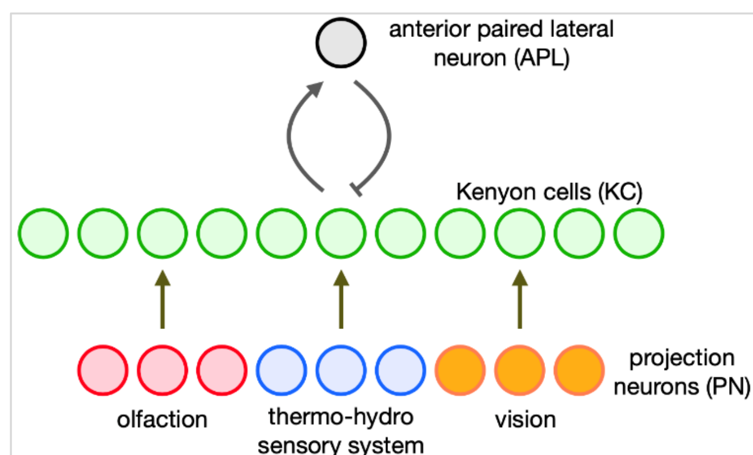


图 3.2 果蝇的网络结构，图片源自[2]。

Yuchen Liang 等人构建了一个类似蘑菇体的神经网络结构[2]，并将其应用于解决自然语言处理，识别文本中单词之间的相关性任务等任务。实验证明，这样的一个网络可以学习到单词的语义表征，并且可以生成静态或者基于上下文的词向量表征。传统的网络结构，如 BERT、GloVe 等，通常使用稠密的词向量表征，而果蝇网络通过训练，可以将单词的语义以及其上下文以一个稀疏的、二值化的哈希码表征。文章同时从词语的相似性度量分析、文本分类以及词义消歧三个方面对学习到的表征进行评价，结果表明，果蝇网络模型不仅可以在一些特定的任务上和一些早期经典的自然语言处理方法相媲美，而且训练和推理消耗的计算资源更小，包括更短的训练时间和更小的内存占用量。这是一个有趣的结果。这项工作除了展示生物脑网络计算的效率，还提出了一些有趣的问题。其中最有意思的问题是：为什么生物神经网络计算是如此的高效？一种解释是，按照进化理论，自然界会选择效率更高的神经网络存留下来。但目前研究者们尚未理解，为什么 Kenyon 细胞形成的网络相比于传统的网络结构具备更高的效率。

3.3 模拟生物视觉系统提高对噪声的鲁棒性

尽管深度神经网络在物体识别任务上取得巨大的成功，但是其对噪声并不具有鲁棒性。比

如在对抗样本问题中，在一张图片上稍微加入一点人眼难以觉察的扰动，就可以让深度网络以很大的置信度做出错误的判断，而人却不受这种噪声的影响。如何构建更加鲁棒的人工智能系统是领域内的一个重大的挑战，而生物的视觉系统为此提供了灵感。

在 2020 年的 NeurIPS 文章上，Joel Dapello 等人在卷积神经网络之前构建了一个模拟初级视觉皮层的模块[3]，发现其可以显著地提高深度卷积神经网络在物体识别时的鲁棒性。简单来说，深度卷积神经网络容易被对抗噪声影响的原因是因为其抽取的视觉特征与灵长目动物视觉系统提取的特征不同。Joel Dapello 等人首先分析了目前主流的深度卷积神经网络，如 AlexNet、VGG、ResNet、DenseNet 以及 MnasNet 等。这些网络均在 ImageNet 数据集上进行了训练，然后用这些网络的第二个隐藏层的特征去解释恒河猴初级视觉皮层（V1, primate visual cortex）的神经活动统计特性。神经元数据集包含了恒河猴 V1 的 102 个神经元，它们的活动通过记录其对 450 个不同的图片刺激（包含了不同的自然纹理以及噪声样本）的反应而得到。研究发现，如果一个卷积神经网络可以更好地解释恒河猴 V1 的神经活动统计特性，那么其对白盒攻击（white box attacks）更加鲁棒。

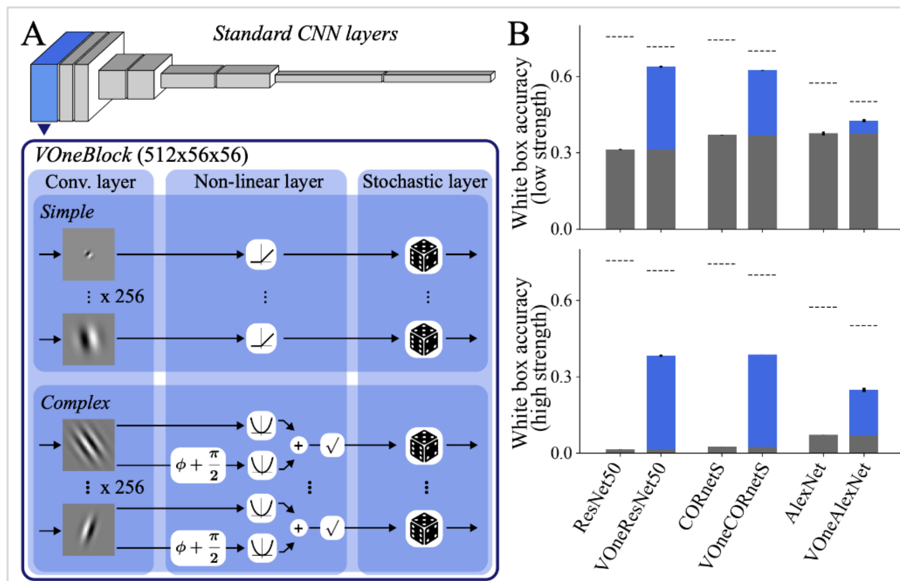


图 3.3 在深度卷积神经网络之前构建一个模拟 V1 的前端网络，组成的系统对白盒攻击更加鲁棒，图片源自[3]。

受此启发，Joel Dapello 等人构建了一个名为 VOneNet 网络架构。这是一个混合网络，由一个初级视觉皮层模型（VOneBlock）和一个传统的卷积神经网络组成，其中 VOneBlock 替代了深度卷积神经网络的第一个隐藏层，如图 3.3 所示。VOneBlock 采用的是计算神经科学中经典的线性-非线性-泊松模型（LNP, linear-nonlinear-Poisson model）。LNP 主要包含固定权值的 Gabor 滤波器组、简单细胞和复杂细胞的非线性操作、以及神经元的随机性。相比于在 ImageNet 数据集上训练的深度卷积神经网络，VOneBlock 可以更好地解释恒河猴 V1 神经元活动的统计特性。实验发现，VOneNet 在 ImageNet 物体识别任务上保持了基准模型的高准确度。同时，作者在一个综合的测试基准上进行了测试，这个测试集涵盖了不同的对抗样本数据和常见噪声的数据。结果表明，VOneNet 表现远超基准模型，甚至比肩一些最先进的噪声防御方法，更重要的是这些优点可以在不同的网络结构间迁移。文章进一步对 VOneBlock 进行分析，发现 VOneBlock 的所有组分均对网络的鲁棒性有贡献，其中神经元响应的随机性在白盒攻击时扮重要的作用，且神经元响应的随机性和 VOneBlock 的特征产生了非线性的交互作用，有效地提升

了网络的鲁棒性。最后，文章显示即使在网络测试阶段去掉神经元的随机性，很大一部分网络的鲁棒性仍然保存，表明在 VOneBlock 训练阶段加入的神经元响应的随机性，可以促使后面的网络层学到更鲁棒的表征。当前的深度卷积神经网络是受到大脑视觉系统的层级化结构的启发而提出的，Joel Dapello 等人的工作进一步表明，如果更为细节地模拟大脑的视觉系统，我们可以在人工智能的视觉任务上收获更多。

3.4 模拟神经群编码机制提高对噪声的鲁棒性

大脑中神经元的群编码原则也可以用于指导构建更为鲁棒的人工神经网络结构。大脑的视觉系统有着海量的神经元，通过神经元群的活动对外部视觉刺激进行表征。这种群编码方式遵循什么样的原则，其背后的编码结构是什么？一种理论认为，大脑群编码遵循的是高效编码原则（efficient coding principle），即神经编码时，会尽可能降低神经元之间对自然图片输入响应的相关性，从而最大化信息的传输，这一编码特点是高维、稀疏，并且信息可以由一个简单的线性网络读出。而另一种理论认为，大脑群编码是将高维度的输入投射到低维度的、平滑的子空间中，在这个子空间里，不同神经元编码的信息具有高相关性，是冗余的，此种编码对噪声更为鲁棒。

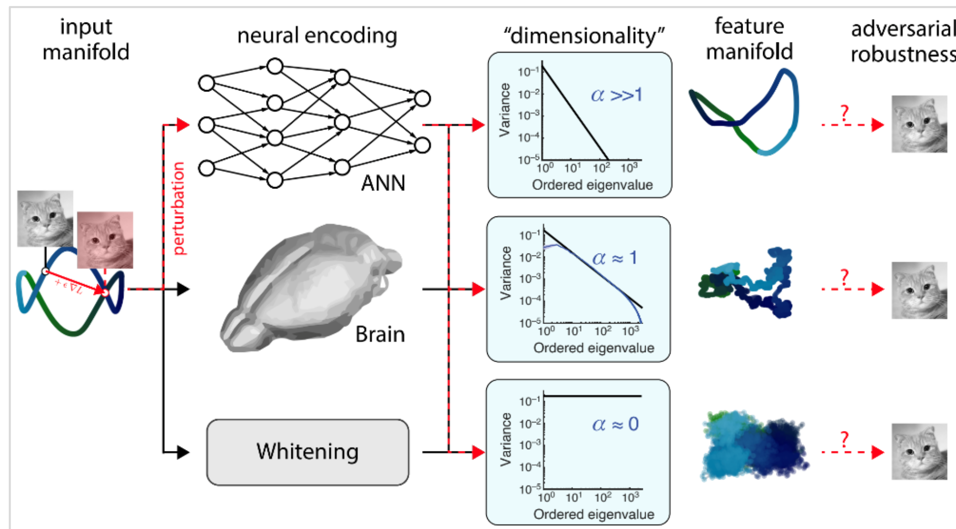


图 3.4 当输入为对抗样本时，不同网络的神经元群协方差矩阵特征值遵循不同的分布，具备不同的鲁棒性，图片源自[5]。

Stringer 等人 2019 年在 Nature 发表了一项研究[4]，通过给小鼠不同的图片刺激，记录小鼠 V1 上万的神经元活动。研究分析发现，大脑 V1 神经元群的编码原则既不是简单地遵循高效性，也不是简单地将高维刺激投射到低维度子空间中，而是在两者之间进行权衡。一方面大脑的初级视觉皮层区的群编码需要维持外部视觉刺激语义之间的相似性，从而促使编码表征具备泛化性能，即保证编码的平滑性；另一方面，在保证编码平滑性的同时，尽可能编码更多的细节信息。文章同时通过理论推导得出，当神经元群编码满足以上条件时，神经元群活动的经验协方差矩阵的特征值分布遵循幂律规律，如图 3.4 所示。当视觉刺激是自然图片时，矩阵的第n个主成分方差按 1/n 缩放。 Josue 等人将此编码特性引入深度神经网络，做了进一步的研究[5]。为了保证网络层对应的特征值分布遵循幂律规律，Josue 等人提出了一种新的正则化方法来确保深度神经网络的神经群活动的协方差矩阵的特征值分布遵循这种幂律规律。他们在 Mnist 数据集上进行测试，发现在对抗样本测试中，被正则项约束的多层感知机和深度卷积神经网络均比基准网络更加鲁棒，而且神经网络的每一层均被正则项约束比只在最后一层施加约

束的效果更好。进一步可视化分析，发现新提出的正则项可以促使网络学习到更具解释性的特征表征。

3.5 总结与展望

近年来，人工智能和神经科学的交流越来越频繁，彼此之间相互启发促进。本章总结了 2020 年几项神经科学对人工智能算法启发的研究工作。总的说来，随着我们对大脑工作机制更加深入地了解，神经科学将会为人工智能发展的思路和算法模型提供更多的帮助。对于未来的发展趋势，我们有如下展望：

（1） 计算神经科学和人工智能的研究对象本质上都是“智能”，前者关注的是生物智能（大脑）的工作原理，而后者则是要发展人工的智能机器。显然，前者的知识应该能为后者发展提供巨大的帮助，但遗憾的是，到目前为止，神经计算的模型和原理在人工智能方面的应用还很少，主要有两方面的原因：

a） 任务的复杂性。目前人工智能的应用主要集中在完成简单的目标分类任务及其各种变种，在这类相对简单、单一的任务上，神经科学的计算法则和模型并不能体现出优势。但当未来人工智能发展走向通用智能，执行更复杂的认知任务，实现在线学习，以及和环境的动态交互时，神经系统的工作原理将会发挥更大的作用。

b） 计算的并行性。神经信息处理的一个基本特色是大规模并行计算，因此计算神经科学发展的原始模型都是包含大量神经元相互作用的微分方程组，这类模型在当前的以冯诺伊曼体系为架构的计算机系统中很难实时实现。目前迅猛发展的神经形态芯片为解决这个问题提供了完美答案。我们预计随着神经形态芯片的成熟，很多在计算神经科学领域已经成熟的计算模型在人工智能领域将得到很好的应用。

（2） 尽管目前还面临一些瓶颈，但如 3.1 和 3.2 节的工作所示，神经系统的网络结构，尤其是进化选择的先天结构，已经开始为人工智能的设计提供了有价值的借鉴。随着各国脑计划的开展，越来越多的神经系统的精细结构、神经活动规律、及其计算功能被解析，这些知识都将加速受脑启发的人工智能发展。

（3） 在具体的模型启示方面，我们觉得如下两个方向有很好的前景：

a） 3.3 和 3.4 节的工作已经提示我们，引入更多生物视觉系统的结构可以进一步提高深度学习网络的表现。在未来一段时间，生物视觉系统的精细结构和工作原理仍将是人工智能受脑启发的重要来源，其中尤其重要的是视觉系统的皮层下结构和计算功能；后者主导了生物视觉的本能行为，在进化上保守，是当前脑科学的研究热点。

b） 最近，神经科学的实验和理论工作都揭示内嗅皮层和海马区的神经元不仅在物理空间的认知地图构建上起到了重要作用，而且它们在抽象概念空间的表征形成和推理过程中也扮演的重要角色。这个重要发现及其计算建模有可能帮助我们阐明人工智能中的一个核心问题，即抽象概念如何和具体数据融合。

参考文献

[1] Lechner, M., Hasani, R., Amini, A., Henzinger, T. A., Rus, D., & Grosu, R. (2020). Neural circuit policies enabling auditable autonomy. *Nature Machine Intelligence*, *2*(10), 642-652.

[2] Liang, Y., Ryali, C. K., Hoover, B., Grinberg, L., Navlakha, S., Zaki, M. J., & Krotov, D. (2021). Can a Fruit Fly Learn Word Embeddings?. arXiv preprint arXiv:2101.06887.

[3] Dapello, J., Marques, T., Schrimpf, M., Geiger, F., Cox, D. D., & DiCarlo, J. J. (2020). Simulating a primary visual cortex at the front of CNNs improves robustness to image perturbations. *BioRxiv*.

[4] Stringer, C., Pachitariu, M., Steinmetz, N., Carandini, M., & Harris, K. D. (2019). High-dimensional geometry of population responses in visual cortex. *Nature*, *571*(7765), 361-365.

[5] Nassar, J., Sokol, P. A., Chung, S., Harris, K. D., & Park, I. M. (2020). On 1/n neural representation and robustness. *arXiv preprint arXiv:2012.04729*.

第四部分 人工智能对神经科学的启发

神经科学中正在涌现的海量数据给我们逐渐带来了对神经系统的全景式的视角，然而相关的理论框架还未建立，如何从海量数据中总结出规律是目前神经科学面对的一个主要挑战。传统的神经科学理论借鉴信息论、系统论、线性矩阵理论等研究方法，对单个神经元对外界刺激的编码和计算特性进行了详细的刻画，并在解释由单个神经元组成的小型网络中的简单计算中取得了成功。但在解释大脑皮层如何实现复杂的认知任务时，记录到的神经元的表征通常和外界任务没有简单的关系，因此我们需要建立更复杂的非线性模型来解释相关数据。

最近的人工智能技术首先为分析处理大规模神经数据提供了新的方法，帮助我们理解在全脑层次上的神经编码原理。其次，深度学习技术的快速发展使得我们能够建立大规模的网络模型来完成复杂的任务，通过比较人工神经网络中的表征方式和大脑中神经元的表征方式，可以为大脑中不同脑区具体完成的计算功能做出预测。通过研究在相关网络模型的训练目标、训练方式、不同正则化规则下得到的表征方式和大脑表征方式的异同，以及比较学习中表征逐步形成的过程，可以帮助理解大脑计算的原理。这可以使神经科学家在投入精力进行相关实验之前，快速检验一系列计算假设，并对相关实验结果做出预测。由于对大脑中的连接结构的直接测量仍然十分困难，而网络模型中所有连接的权重很容易得到，通过分析相关神经网络中的连接结构和动力学特点，可以对相关计算功能在具体神经环路层面的实现方法提供参考。

更有意思的是，针对目前人工神经网路模型的弱点的思考可以指导设计新颖的神经科学实验，为神经科学研究提供新的角度。另外，深度学习的成功凸显了基于反向传播算法的有监督学习的重要性，而由传统的神经突触可塑性理论启发的生物可信的学习算法虽然在解释无监督统计学习领域取得了成功，但在有监督学习任务上的效果还和反向传播算法有差距。受到反向传播算法启发和挑战，近期如何结合局部的神经可塑性和大脑中存在的神经环路结构去完成复杂的学习任务是计算神经科学的一个活跃领域，也启发了相关实验来寻找可能在相关学习过程中起关键作用的神经元类型和结构。

接下来，我们将从深度强化学习理论如何帮助理解大脑多巴胺神经元的编码、类似反向传播算法的学习可能如何在大脑中实现、结合海马结构和贝叶斯学习对海马建模三个具体的例子，来总结 2020 年取得的一些有重要启发性的科研成果。

4.1 强化学习理论帮助理解大脑的奖惩机制

脑科学研究的一个重要的问题是：外部的奖赏信息如何引导学习、决策、以及神经表征的形成？深度强化学习将深度神经网络和经典的强化学习整合，为研究智能体如何实现学习，表征和决策之间的交互，提供了一个全面的框架。近年来，深度强化学习在围棋博弈，预测蛋白质折叠等任务上取得了巨大的成功。那么深度强化学习框架是否可以帮助我们更加深入地理解大脑中的强化学习机制？

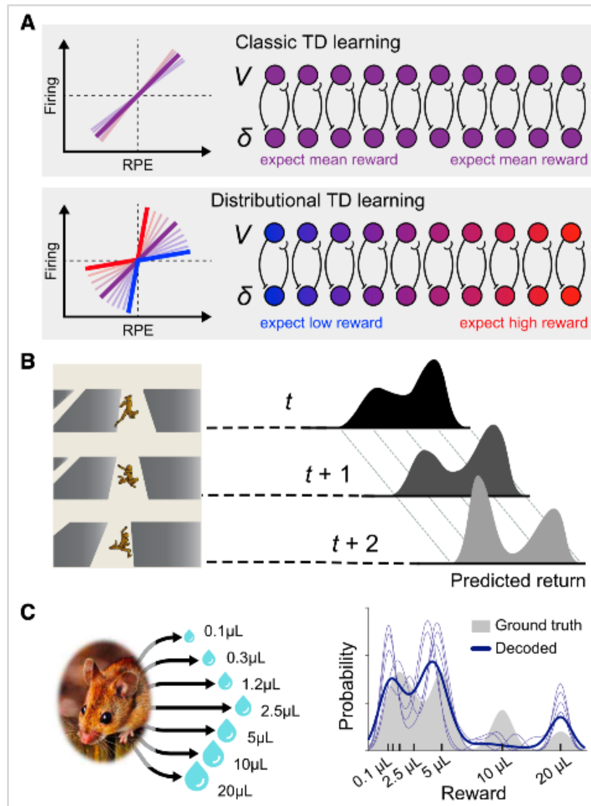


图 4.1 分布式强化学习，图片源于[2]。

受启发于人工智能中的分布式强化学习算法[1]，Will Dabney 等人在 2020 年的 Nature 上发文指出[2]，大脑中可能存在分布式强化学习系统，大脑不是以一个平均值来预期未来可能的回报，而是以一个“概率分布”来进行预期。具体来说，大脑中的多巴胺表征了一种“惊讶”信号，当结果好于预期时，多巴胺会被更多的释放。在经典的强化学习理论中，多巴胺神经元的活动被认为是相同的，它们的活动编码了一个简单的标量，表征一个预期回报误差。如果标量为正值，则表征一个积极的惊喜信号，如果为负值，则表征一种对结果的失望。但实验中观测到多巴胺神经元的放电经常具有多样性。越来越多的神经计算模型研究表明，考虑回报预期的分布和不确定性对于在一个不确定环境中进行决策非常重要。在人工智能的分布式强化学习算法中，回报预测误差是一个向量，而非一个标量，向量中的不同值表征不同先验预测下的回报预测误差，涵盖了从高度乐观的预测到高度悲观预测范围。人工智能模型实验证明以上细节可以显著地提高模型在不同任务中的表现[1]。受此启发，Dabney 等人训练老鼠完成一个复杂任务，任务回报是一个分布，即小鼠无法精确地预测其得到的奖励。作者通过分析研究了小鼠多巴胺神经元的活动记录，发现不同多巴胺神经元表征的值不同，并形成一个分布，其包含了一个预测误差信号的分布，从乐观到悲观的预期，如图 4.1 所示。

4.2 反向传播算法在大脑可能实现的机制

理解大脑的学习机制一直是神经科学的经典研究问题。著名的加拿大生理心理学家赫布早在 1949 年就指出，神经网络的学习发生于神经元之间的突触，而突触连接强度的改变依赖于突触前后神经元之间的活动；当两个神经元同时发放，两者之间的神经元连接强度增强。在赫布理论的指导下，神经科学研究发现了各种神经元连接权重的局部更新法则，但单个神经元突触的变化如何与复杂的网络行为或者功能构建起联系起来一直是个谜。在人工智能领域，反向传播算法被用来训练深度网络，是深度学习技术成功的关键，研究人员也开始重新思考大脑是否可以实现类似反向传播的学习算法。生物大脑实现反向传播算法遇到的主要问题为（如图 4.2 所示）：第一，经典的反向传播算法分为两个阶段，前向传播阶段和反向传播阶段，反向传播过程需要利用到前向传播过程得到的信息。计算机很容易协调并实现上述过程，但生物神经网络很困难；第二，突触权值转置难题，即生物神经网络可以通过反馈连接实现反向传播过程，但为了计算出神经元梯度更新值，需要反馈连接权重是前馈连接突触权值转置，但生物上每一个神经元仅仅接受与其有连接的神经元的发放活动信息，而不知道其它神经元的突触连接信息，即大脑的突触权重更新是一个局部计算过程。

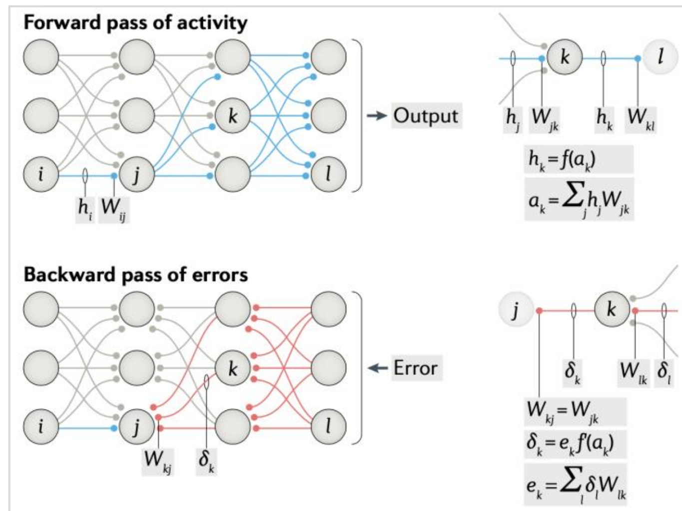


图 4.2 反向传播算法，图片源于[3]。

近年来，不少理论研究试图解决反向传播算法的生物学合理性问题。为了解决突触权值转置难题，Timothy P. Lillicrap 等人在 2016 年提出反馈校准（feedback alignment）算法[4]，即大脑的反馈连接并不需要和前馈连接权值的转置保持精准的一致，一个随机的、固定不变的反馈连接矩阵，也可以帮助神经网络估算有效的梯度信息，并沿着目标函数下降的方向更新突触权值。在学习过程中，前馈连接权重会逐步和反馈连接权重对齐，从而帮助网络估算有效的梯度信息。由于反馈校准算法对梯度的更新并不精准，故用此算法训练网络需要大量的数据以及花费较长的时间获得收敛。Bengio 等人提出稳态传播（equilibrium propagation）算法[5]，由于存在互馈和反馈连接，网络在外部输入的驱动下，神经元响应会在网络中传播回荡，最终达到一个稳态 A，并得到一个网络输出。接着再给予网络输出神经元一个输入驱动，迫使输出神经元的活动和训练目标值一致，网络在此驱动下会达到一个新的稳态 B。最终基于以上两种稳态情形，可以获取计算梯度的相关信息，并基于对比赫布学习法则更新网络。训练网络时，只需要在标注数据上不断地重复稳态传播过程即可。总的讲，反向传播算法为我们理解大脑如何学习提供了一个概念框架，在网络的功能层面和局部生物神经元的学习法则之间构建起一座桥梁。

2020 年，Timothy P. Lillicrap 等人对大量相关研究进行了梳理[3]，并总结出了一个可能生物合理性框架，相关文章发表在 Nature Reviews Neurosciences 上。为了实现反向传播算法，大脑需要计算、表征和传播梯度信息，那么生物系统如何实现上述过程？Timothy P. Lillicrap 等人总结并指出，大脑中的神经网络可以通过神经元活动之间的差值信号来表征神经元的梯度信息（NGRAD，neural gradient representation by activity differences），神经元之间的突触权值基于神经活动的差值进行更新。在 NGRAD 框架假设下，生物神经网络不需要传播误差梯度，而只需要传播神经活动信息。NGRAD 实现类反向传播算法的核心观点是，高层次的神经元活动可以推动低层次的神经元活动到某一个水平，从而更好地和高层次活动或者某些目标保持一致性，这种高层次活动可能来自于一个目标，不同的神经模态，或者较大尺度的空间和时间上下文。这种推动作用会导致低层次神经元的活动产生变化。这种变化结合局部的神经活动，可以被用于计算类似反向传播算法的突触权重更新值。NGRAD 本质上是通过自上而下的神经活动来驱动网络的学习，可以基于大脑中大量的反馈连接实现。目前基于 NGRAD 只是提供了一种的高层次观点，而真实的神经环路如何实现类似梯度反向传播的学习算法，未来还需要更多的研究。比如前馈通路和反馈通路如何实现协同作用，如何计算神经元活动模式之间的差异，然后如何利用这些计算出的差异值进行合适的突触权值更新，对于以上问题还不完

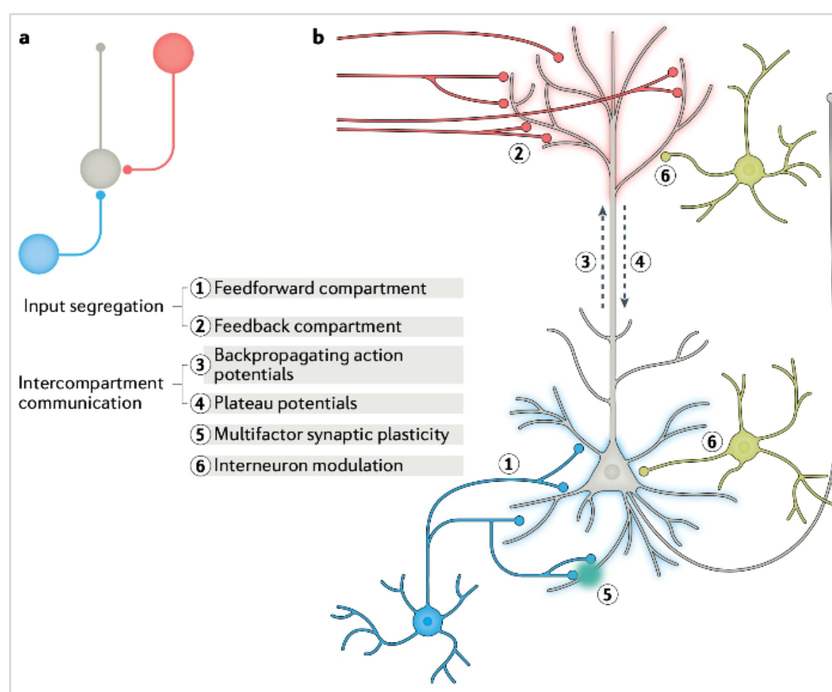


图 4.3 实验发现提供的一种可能性，大脑可能如何实现类反向传播的学习算法，图片源自[3]。

全清楚，图 4.3 提供了一种可能的生物物理实现。反向传播算法假设网络存在一个精准的监督信号，在大脑中，这个监督信号或许是由一团高层级的神经元群计算得出。但是真实的神经系统中，每一个神经元并不总是有精准的监督信号，而是依赖于环境反馈的奖赏或者惩罚信息，例如运动皮层区，网络只会基于反馈连接获取一个粗略的强化信号，此时大脑如何实现层级化网络的优化学习呢？Pieter R. Roelfsema 实验组结合注意力机制和全局误差信号给出了一个算法框架 BrainProp（Attention-Gated Brain Propagation）[6]，相关文章发表在 2020 年的 Neurips 上。早在 1990 年，Roelfsenma 和其合作者即发现，当猴子盯着一个物体时，皮层中表征这个物体的神经元会变得很活跃，即猴子集中注意力盯着物体会产生一个反馈信号，作用于任务相关神经元。此时反馈连接介导的信号并不是一个误差信号。另外，当动物选择一个

动作，产生的回报高于预期时，大脑中的多巴胺系统被激活，释放神经调质作用于相关区域。此时多巴胺神经元活动扮演着一个全局的强化信号，通过反馈连接作用于并调控神经元的可塑性强度，使那些和强化信号相关的神经元突触可以更好地被局部赫布学习法则更改。实验发现基于注意力的 BrainProp 算法在多个数据集上，如 CIFAR10，Tiny ImageNet 等，性能优于现有的一些生物合理性的算法。

4.3 神经网络建模帮助理解大脑海马中信息表征的机制

理解大脑的工作机制，我们需要理解大脑中的神经表征是如何产生的，以及其如何帮助实现相关的认知行为的。大量的研究发现，大脑海马区的位置细胞和内侧内嗅皮层的网格细胞在空间表征和空间导航中扮演着重要的角色。那么这些神经表征是如何产生的？负责空间导航的神经环路是否也可以被大脑用于非空间的，抽象的概念关联结构推理？如果可以，大脑会如何实现这一点？在 2020 的 Cell 文章中，Whittington 等提出了一个内嗅皮层-海马区计算模型[7]，模型实验研究表明，其可能在大脑进行抽象的，不同类型的关联结构推理中扮演着重要的角色。这一模型研究工作受启发于 Tolman 和 Eichenbaum 等人的工作，为致敬他们的杰出研究，模型被命名为 Tolman-Eichenbaum 模型（TEM）。Tolman 在 1948 年首先提出了认知地图的概念，Tolman 认为大脑中存在认知地图，可以用于空间导航。其后这一概念被研究者泛化，大脑中的认知地图既可以表征物理空间，也可以表征抽象概念空间，帮助实现编码，存储，提取空间中的相对位置关系和对应的物体属性等，从而实现物理空间以及抽象概念空间相关结构的推理。Eichenbaum 等人进一步指出了大脑海马区在物理空间表征和抽象概念空间表征中的作用。

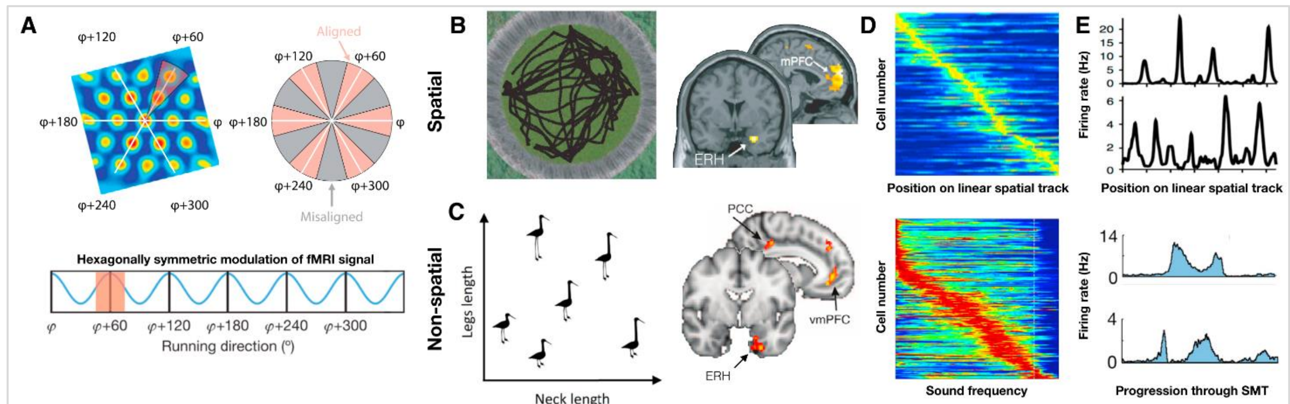


图 4.4 位置细胞和网格细胞既可以表征二维的物理空间，也可以表征非物理空间，如抽象的概念空间[8]。

近年来的很多实验研究支持以上观点。比如在 Constantinescu 等人的研究中[9]，受试者观看连续渐变的卡通鸟，卡通鸟的外观会在两个维度上变化，即变化脖子和腿的长度，从而形成一个二维的概念空间，如图 4.4B 所示。受试者被要求预测并报告什么时候呈现的刺激会和某个目标小鸟的外观一致，然后给予相应的奖赏。连续渐变的刺激在二维概念空间上，可以表征为一个向量。通过对 fMRI 的活动进行分析，发现当受试者在该鸟空间中导航时，内嗅皮层中也会出现类似网格细胞的表征，而海马区会出现对特定的抽象刺激响应的神经元，类似于空间导航中的位置细胞。这表明大脑的内嗅皮层可以表征一个二维的抽象概念空间，每一个维度可以是某个概念的特征。在另一项对老鼠海马区的研究中[10]，让老鼠抓住一个杠杆，同时增加一个音调的频率。当频率增加至一个目标频率时，老鼠需要释放杠杆，从而获得奖励。尽管这个任务不是一个空间性质的任务，但是却和一维空间任务具备类似的拓扑结构，音调频率从

低到高类似于在线性轨道中从一端到到另一端的过程。此时海马区和内嗅皮层区会分别出现类似位置细胞的和网格细胞的表征，只不过编码的不是空间位置，而是音调频率，如图 4.4D 所示。而在非人灵长类，以及灵长类动物中，研究也发现网格细胞还可以编码眼睛在二维图片上的注视点位置。以上实验均表明，海马及其相关区域，并不仅仅用于表征二维的物理空间，也可以表征非物理空间，如在人的实验中表征抽象的概念空间。

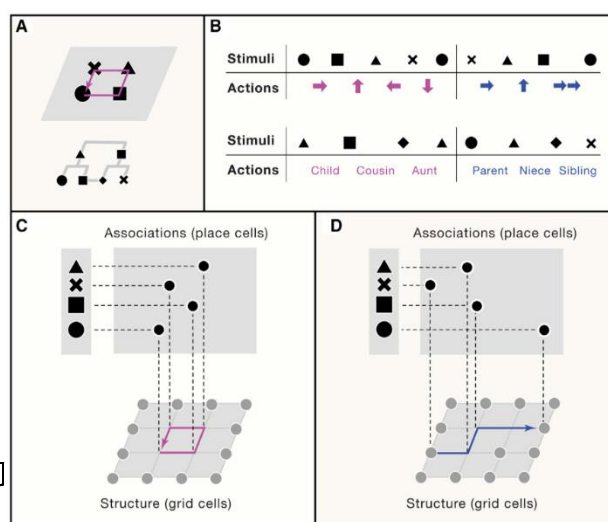


图 4.5 网格细胞和位置细胞组成的系统可以实现不同环境下结构关系的迁移，图片源自[11]。

基于以上实验，Whittington 等人在理论模型中假设，网格细胞编码一个抽象的空间，而位置细胞则编码特定的情景记忆，外界的感知输入和抽象的空间位置可以通过联合编码的方式存储在位置细胞的互馈连接中。例如当编码物理空间时，网格细胞会整合当前的运动向量来产生一个更新后的位置估计，如图 5A 所示。而位置细胞为构建各种输入之间的记忆关联，包括来自网格细胞的状态以及来自外界地标的视觉表征，如图 4.5C 所示。如果给一个外部的地标输入，如图 4.5C 中的圆圈，基于位置细胞的关联记忆，可以提取出相应的网格细胞状态，其表征了圆圈的相对位置关系，如物理空间的左下角。类似地，给一个网格细胞状态，可以由关联记忆激活相应的地标视觉表征。当我们在大脑表征的物理空间中“旅行”时，我们可以基于以前的关联记忆，预测出相应的地标，这种预测甚至可以沿着一个我们从未体验过的路径进行，这种能力也被称之为零样本推理。研究发现，大脑这一种零样本推理能力不仅仅局限于物理空间，也可以被用于抽象的概念空间。

在 2020 的 Cell 中，Whittington 等人采纳了以上的理论假设[7]，并构建了一个可以工作的计算模型，模型训练基于人工智能领域发展的一种变分推断的学习框架，实验发现其可以在具备相似结构的环境中，实现高效地推理。TEM 是一个结构化的模型，包含两层互馈神经网络，一个网络表示内嗅皮层，一个网络表示海马区，模型假设外部的地标输入和动作输入是解耦的，内嗅皮层网络接受外部的动作输入，学习如何基于动作输入转化为相应的空间中的位置，而海马网络接受外部地标输入，通过联结编码学习外部地标输入位于哪个空间位置上，训练模型实现零样本的地标预测任务。实现发现：（1）模型可以实现空间导航任务以及非空间的关系推理任务；（2）在内嗅皮层网络和海马区网络形成多种神经元的表征，和实验观测一致。这项研究为大脑中海马区和内嗅皮层区的多种神经元的表征提供了一个统一的框架解释。这一研究模拟了网格细胞和海马细胞的结构关系，通过将网络解离为共享动作的转移模块和一个联想记忆模快，解决了新环境中的快速适应和快速推理任务。这项研究向我们展示了大脑可能如何学习和理解抽象的关系表征。

4.4 总结与展望

近年来，人工智能和神经科学的交流越来越频繁，彼此之间相互启发促进。在本文中，我们从理解大脑的角度出发，围绕几个重要的问题，对 2020 的相关工作进行了总结。总体说来，由于神经科学中新技术的发展正在产生海量的数据，借助人工智能算法建模和理解大脑的工作机制变得越来越重要，我们结合本文总结的研究进展对未来总的发展趋势做出如下展望：

（1）比较大脑和人工神经网络中的信息表征方式是一个目前的重要趋势。具体来说：

大脑的视觉系统比深度神经网络更加鲁棒，可以更好地处理复杂的视觉场景任务，研究在复杂视觉任务中大脑神经元的编码方式是一个重要的方向。最近的深度学习智能模型已不再局限于物体分类，而是试图运用生成式模型对整个场景进行描述和理解，这也许更接近大脑的计算目标，对比相关模型和大脑的表征是一个有前景的方向；

人工智能技术为我们提供了一种自上而下的方式，研究神经元的表征形成。神经元的表征形成由网络结构和目标任务共同作用产生。由于大脑处理任务需要多个脑区协同实现，结合大脑的连接结构及编码方式和深度学习算法来训练模块化的多脑区网络模型，可以帮助我们更好地理解不同皮层区域的表征是如何形成的，并对其在信息处理过程中各环节的分工合作给出一个统一的解释。

（2）大脑的学习功能具有目前人工神经网络不具备的灵活性、自主性等优势。然而尽管目前有众多生物合理性的算法被提出可以用于优化大脑中的多层级网络结构，但这些算法通常只被用于解决简单的任务，在复杂的任务场景上，这些生物合理性的算法与标准的反向传播算法还存在较大差距。NGRAD 等理论框架尝试解决生物大脑如何实现多层级网络的优化，主要用于处理静态信息。当生物大脑处理动态时序信息时，如何沿时序的进行信用分配是一个重要研究方向。另外，不需要反向传播的在小数据和实时条件下能实现高效学习的新型学习算法也许和大脑需要解决的问题更相似，这也许能为帮助理解大脑的学习方式提供新的启发[12]。未来一个有意思的方向是发展可以解决复杂任务的生物合理性算法，或许在复杂任务上才能真正体现生物算法的优越性。另外研究人在实现复杂任务时如何协调多种学习机制以及逐步学习的过程也可能为人工智能系统提供启发。

（3）深度强化学习研究智能体如何在复杂环境中获取奖励，这和生物智能需要完成的计算功能很相似。在深度强化学习领域，最近的一个趋势是将对世界的建模和传统强化学习结合来完成有模型的强化学习，尤其是和不同记忆模块的结合，如情景记忆和工作记忆模块等，用于解决一些具有挑战性的认知任务。记忆和决策系统都是神经科学领域的重要研究主题，这些最新的模型为我们研究大脑中记忆系统如何帮助实现基于奖赏的学习和决策提供了一个很好的模型框架。这些模型可以用于引导神经科学的实验设计和研究。神经科学的研究已经发现人类在解决最能体现人的智能的抽象推理、社会认知等任务时，运用到的大脑神经回路和目前深度强化学习在解决简单决策，导航等任务时采用的结构类似。由于相关领域的研究也在起步阶段，将目前的深度强化学习理论运用到解释抽象推理、社会认知等复杂任务中的神经活动也是一个趋势。相关模型还可以被运用到解释精神性疾病中决策、情绪等功能的缺陷，为相关疾病的诊断和治疗提供线索。

参考文献

[1] Bellemare, M. G., Dabney, W., & Munos, R. (2017, July). A distributional perspective on reinforcement learning. In *International Conference on Machine Learning* (pp. 449-458). PMLR.

[2] Dabney, W., Kurth-Nelson, Z., Uchida, N., Starkweather, C. K., Hassabis, D., Munos, R., & Botvinick, M. (2020). A distributional code for value in dopamine-based reinforcement learning. *Nature*, *577*(7792), 671-675.

[3] Lillicrap, T. P., Santoro, A., Marris, L., Akerman, C. J., & Hinton, G. (2020). Backpropagation and the brain. *Nature Reviews Neuroscience*, *21*(6), 335-346.

[4] Lillicrap, T. P., Cownden, D., Tweed, D. B., & Akerman, C. J. (2016). Random synaptic feedback weights support error backpropagation for deep learning. *Nature communications*, *7*(1), 1-10.

[5] Scellier, B., & Bengio, Y. (2017). Equilibrium propagation: Bridging the gap between energy-based models and backpropagation. *Frontiers in computational neuroscience*, *11*, 24.

[6] Pozzi, I., Bohte, S., & Roelfsema, P. (2020). Attention-Gated Brain Propagation: How the brain can implement reward-based error backpropagation. *Advances in Neural Information Processing Systems*, *33*.

[7] Whittington, J. C., Muller, T. H., Mark, S., Chen, G., Barry, C., Burgess, N., & Behrens, T. E. (2020). The Tolman-Eichenbaum machine: Unifying space and relational memory through generalization in the hippocampal formation. *Cell*, *183*(5), 1249-1263.

[8] Behrens, T. E., Muller, T. H., Whittington, J. C., Mark, S., Baram, A. B., Stachenfeld, K. L., & Kurth-Nelson, Z. (2018). What is a cognitive map? Organizing knowledge for flexible behavior. *Neuron*, *100*(2), 490-509.

[9] Constantinescu, A. O., O'Reilly, J. X., & Behrens, T. E. (2016). Organizing conceptual knowledge in humans with a gridlike code. *Science*, *352*(6292), 1464-1468.

[10] Sanders, H., Wilson, M., Klukas, M., Sharma, S., & Fiete, I. (2020). Efficient Inference in Structured Spaces. *Cell*, *183*(5), 1147-1148.

[11] Aronov, D., Nevers, R., & Tank, D. W. (2017). Mapping of a non-spatial dimension by the hippocampal–entorhinal circuit. *Nature*, *543*(7647), 719-722.

[12] Budden，D., Marblestone, A., Sezener,E.,Lattimore, T., Wayne, G., Veness, J. (2020). Gaussian Gated Linear Networks. arXiv:2006.05964v2 [cs.LG] https://arxiv.org/abs/2006.05964

第五部分 结语

1979 年，在美国 Kansas City 机场附近的一个宾馆里，来自语言学、心理学、计算机科学、神经科学、哲学和人类学的 11 名学者汇聚在一起，成立了认知科学学会。直到这一刻，一个以揭秘心智（Mind）奥秘为目标的新兴学科——认知科学——以一个独立的学科出现在世人面前。由于认知科学是多个学科融合的结果，因此必然具有不同的方法论和术语，但是这个学科的所有研究者都坚信：“认知即信息加工，而信息加工即计算”。

这个假设的来源可以追溯到上世纪 40 年代末 50 年代初。在当时，有两个在科学技术史上重大的事件发生：（1）在实践层面，逻辑学、数学和电子工程学的深度交叉孕育出了数字计算机，它为人类心智提供了一个类比的实体（Metaphor）；（2）在理论层面，一门关于智能体适应性行为（Adaptive Behaviors of Intelligent Agents）的新科学——控制论（Cybernetics）——诞生了，该学科从一开始就致力于将机器智能与生物智能整合在一起以探索思维的本质。这一传统，从控制论大家钱学森在晚年致力于思维科学的研究就可见一斑。因此，一般认为，控制论就是认知科学的前身。

在类比实体和理论的双重加持下，形成了以符号主义（Symbolism）为核心思想的认知科学。在这里，计算即基于规则的符号操纵（Rule-Governed Symbol Manipulations）。符号主义认知科学在 50 年代到 80 年代里取得了一系列耀眼的成果：例如，数学定理的形式化证明、计算语言学、以国际象棋为核心的问题解决、专家系统等，甚至图灵奖得主 Allen Newell 乐观地提出了关于认知的大一统理论（Unified Theories of Cognition）。

但是，在符号主义认知科学的上空始终有几片挥之不去的乌云，例如语音识别、语言翻译、客体识别等等。在契而不舍的努力之后，符号主义认知科学家不得不承认其局限性，即符号主义对于那些有清楚定义的知识和目标态的问题（well-posed problems）能提供较好的解决方案，而对于那些不适定的问题（ill-posed problems）则困难重重。遗憾的是，不适定的问题远远多于适定的问题。认知哲学家 Clark 反思到：“我认为关于心智的任何研究不能在生物的真空中开展”。也就是说，关于心智的研究不能离开大脑这个生物基础而独立存在。于是，受生物神经网络启发的连接主义认知科学应运而生。

事实上，连接主义认知科学是与符号主义认知科学几乎同时出现。在上世纪 40 年代美国心理学家 McCulloch 和数学逻辑学家 Pitts 模仿神经元的结构和工作原理提出了一个神经元的数学模型（MP 模型）；在 50 年代，心理学家 Rosenblatt 在此基础上提出了被称为感知机（Perceptron）的像人类心智一样“可感觉、可识别、可记忆和可反应”的第一个人工神经网络在 80 年代，心理学家 Rumelhart 和 McClelland 参照生物神经网络提出了并行分布式加工（Parallel Distributed Processing, PDP）理论，正式宣告连接主义（Connectionism）认知科学的诞生。连接主义认知科学抛弃了符号与中央控制等符号主义认知科学的核心，强调智能来源于简单处理器的并行分布式加工。这一理念，为后来的深度学习的繁荣提供了理论基础。近年来，深度学习在封闭环境和静态规则领域取得了巨大的进步，但是面对开放环境和动态规则，深度学习又面临新的挑战。面对这一困境，认知科学家正试图将心智、大脑与环境统一起来，提出了具身主义认知科学（Embodied Cognition）。这正如心理学家 Gibson 所说：“我们经常忽略一个事实，即动物和与其所处的环境是密不可分的”。

从上世纪 40 年代开始，认知科学从萌芽到壮大，从分歧到深入，一方面我们看到的是揭秘心智任重而道远；在另外一方面，虽久经磨难，认知科学也在不断的自我革命，自我成长。而这成长的第一推动力，就是学科交叉。从控制论到神经科学到心理学，认知科学不断地获得

新的动力。

在2020年9月，来自系统神经科学、认知神经科学、计算神经科学的六位学者和一群年轻的博士后在北京智源人工智能研究院聚集一起，期望利用“人工智能的认知神经基础”重大研究方向的多学科交叉基础，为探索心智的奥秘做一点原创的贡献。这本仓促而就的白皮书总结了过去一年间在这些交叉前沿领域的一些重要探索和发现，以期更好的为相关领域的研究者搭建沟通的平台和桥梁。这是我们的第一次尝试，因水平有限、时间仓促，难免挂一漏万，还请读者原谅。但是，我们坚信，这小小的一步，会是组成我国认知科学将来大步向前不可或缺的组成部分。

关于我们

智源“人工智能的认知神经基础”重大研究方向

智源“人工智能的认知神经基础”重大方向(Brain and Machine Intelligence)旨在将神经科学、认知科学和信息科学进行交叉融合，加强人工智能和脑科学的双向互动和螺旋发展，揭示生物智能系统的精细结构和工作机理，构建功能类脑、性能超脑的智能系统，以视觉等功能和典型模式动物作为参照物构建智能水平测试平台，为人工智能未来发展探索可行道路。

北京智源人工智能研究院

北京智源人工智能研究院（Beijing Academy of Artificial Intelligence，简称BAAI）成立于2018年11月，是在科技部和北京市委市政府的指导和支持下，由北京市科委和海淀区政府推动成立的新型研发机构。

智源研究院的愿景和目标是按照国家新一代人工智能发展规划总体部署，聚焦原始创新和核心技术，建立自由探索与目标导向相结合的科研体制。支持科学家勇闯人工智能科技前沿“无人区”，挑战最基础的问题和最关键的难题，推动人工智能理论、方法、工具、系统和应用取得变革性、颠覆性突破。营造全球最佳的学术和技术创新生态，推动北京成为全球人工智能学术思想、基础理论、顶尖人才、企业创新和发展政策的源头，率先成为国际领先的人工智能创新中心。推动人工智能产业发展和深度应用，改变人类社会生活，促进人类、环境和智能的可持续发展。

编者介绍

陈路瑶 智源博士后（人工智能的认知神经基础重大研究方向），华中科技大学博士，指导老师是神经接口与康复领域专家何际平教授和集成电路与微电子领域专家邹雪城教授，海外加州大学洛杉矶分校联合培养博士，主要研究方向为神经调控。曾在 Journal of Neural Engineering、Sensors、IEEE Sensors Journal 、IEEE EMBS on Neural Engineering、IEEE CBS 等期刊与国际会议发表论文 7 篇，曾参与多项国家级项目，主持一项校创新基金。

刘　祥 智源博士后（人工智能的认知神经基础重大研究方向），北京大学理学博士，主要研究方向为系统神经科学数据分析与建模。

徐琳璐 智源博士后（人工智能的认知神经基础重大研究方向），中国科学院大学生物物理研究所理学博士，主要研究方向为斑马鱼学习记忆的神经机制、斑马鱼视知觉。

邹晓龙 智源博士后（人工智能的认知神经基础重大研究方向），北京师范大学博士。主要研究方向为计算神经科学，类脑计算。

联系方式

闫亚琼　科技创新部

邮箱：yqyan@baai.ac.cn

地址：北京市海淀区清华科技园赛尔大厦 5 楼



添加智源小助理微信，
加入“BAAI 认知神经基础深度交流群”



智源社区